THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY

CARL MURCHISON

JOHN E. ANDERSON
University of Minnesota

CHARLOTTE BÜHLER
Universität, Wien

CYRIL BURT
University of London

LEONARD CARMICHAEL
Tufts College

SANTE DE SANCTIS
R. Università di Roma

ARNOLD GESELL
Yale University

WILLIAM HEALY
Judge Baker Foundation, Boston

BUFORD JOHNSON
The Johns Hopkins University

HAROLD E. JONES
University of California

TRUMAN L. KELLEY
Harvard University

K. S. LASHLEY
Harvard University

A. R. LURIA
Medico-biological Institute, Moscow

HENRI PIÉRON
Université de Paris

GEORGE D. STODDARD
New York State Department of Education

CALVIN P. STONE
Stanford University

LEWIS M. TERMAN
Stanford University

GODFREY THOMSON
University of Edinburgh

E. L. THORNDIKE
Teachers College, Columbia University

C. J. WARDEN
Columbia University

JOHN B. WATSON
New York City

HELEN THOMPSON WOOLLEY
Teachers College, Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 65

1944

Copyright, 1944, by The Journal Press

Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Adelsheim, Elizabeth 97
Ames, Louise Bates 219
Arnold, Martha 291
Baer, J. A. 53
Bijou, Sidney W. 133, 147
Bradway, Katherine P. 197
Brown, Fred 161
Child, Irvin L. 97
Dittmann, Allen T. 265
Graham, V. 227
Jackson, Joseph 113
Jackson, T. A. 227
Klugman, Samuel F. 89
Kuhlen, Raymond G 291
Leeper, Robert 3
Lewis, W. Drayton 183
Long, L. 227
McCandless, B. R. 147
Morgan, John J. B. 281
Morgan, Sarah S. 281
Pegram, Edna Lee 239
Reymert, Martin L. 301
Robins, Margaret E. 177
Schilder, Paul 67
Scow, Robert O. 177
Spoerl, Dorothy Tilden 307
Stephens, J. M. 53
Thorne, Frederick C 275
Welch, L. 227
Wellman, Beth L. 239

TABLE OF CONTENTS

Dr. Hull's *Principles of Behavior* 3
By Robert Leeper

Factors influencing the efficacy of punishment and reward: The opportunity for immediate review, and special instructions regarding the expected rôle of punishment 53
By J. M. Stephens and J. A. Baer

Congenital alexia and its relation to optic perception 67
By Paul Schilder

Test scores for clerical aptitude and interests before and after a year of schooling 89
By Samuel F. Klugman

The motivational value of barriers for young children 97
By Irvin L. Child and Elizabeth Adelsheim

A survey of psychological, social, and environmental differences between advanced and retarded readers 113
By Joseph Jackson

Behavior efficiency as a determining factor in the social adjustment of mentally retarded young men 133
By Sidney W. Bijou

An approach to a more comprehensive analysis of mentally retarded predelinquent boys 147
By S. W. Bijou and B. R. McCandless

An experimental and critical study of the intelligence of Negro and White kindergarten children 161
By Fred Brown

SHORT ARTICLES AND NOTES 177

Apparatus for the study of delayed reaction in children (Robert O. Scow and Margaret E. Robins) 177

BOOKS 181

CRITICAL REVIEWS OF RECENT BOOKS 183

L. S. Hollingworth. *Children Above 180 IQ*. (Reviewed by W. Drayton Lewis) 183

BOOKS RECENTLY RECEIVED 193

IQ constancy on the revised Stanford-Binet from the preschool to the junior high school level 197
By Katherine P. Bradway

Early individual differences in visual and motor behavior patterns: A comparative study of two normal infants by the method of cinemanalysis 219
By Louise Bates Ames

Generalization of the concept of middleness 227
By V. Graham, T. A. Jackson, L. Long, and L. Welch

Binet IQ changes of orphanage preschool children: A re-analysis . . 239
By Beth L. Wellman and Edna Lee Pegram

A preliminary study of audition in the white rat 265
By Allen T. Dittmann

The inheritance of shyness in dogs 275
By Frederick C. Thorne, M D.

Infant learning as a developmental index 281
By John J. B. Morgan and Sarah S. Morgan

Age differences in religious beliefs and problems during adolescence . . 291
By Raymond G. Kuhlen and Martha Arnold

SHORT ARTICLES AND NOTES 301

An apparent case of neurotic influence on somatic function (Martin L. Reymert) 301

BOOKS 301

CRITICAL REVIEWS OF RECENT BOOKS 307

A. Gesell and F. L. Ilg. *Infant and Child in the Culture of Today.* (Reviewed by Dorothy Tilden Spoerl) 307

BOOKS RECENTLY RECEIVED 315

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

September, 1944
Volume 65, First Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND
JOURNAL OF GENETIC PSYCHOLOGY

Psychological Fou

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

John E. Anderson
University of Minnesota

Charlotte Bühler
Universität, Wien

Cyril Burt
University of London

Leonard Carmichael
Tufts College

Sante De Sanctis
R. Università di Roma

Arnold Gesell
Yale University

William Healy
Judge Baker Foundation, Boston

Buford Johnson
The Johns Hopkins University

Harold E. Jones
University of California

Truman L. Kelley
Harvard University

K. S. Lashley
Harvard University

A. R. Luria
Medico-biological Institute, Moscow

Henri Piéron
Université de Paris

George D. Stoddard
New York State Department of Education

Calvin P. Stone
Stanford University

Lewis M. Terman
Stanford University

Godfrey Thomson
University of Edinburgh

E. L. Thorndike
Teachers College, Columbia University

C. J. Warden
Columbia University

John B. Watson
New York City

Helen Thompson Woolley
Teachers College, Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special Edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

Copyright 1944, by The Journal Press
Provincetown, Massachusetts
Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879
Reëntered as second-class matter, May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publications shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price will be $7 per volume.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinco or halftone cuts.

(6). The publication charges will be as follows:

4 page signature	$18
8 " "	24
12 " "	36
16 " "	48
20 " "	60
24 " "	72
28 " "	84
32 " "	96
36 " "	108
40 " "	120

These charges are tentative and will be adjusted as justified by the subscription list (especially by the effect of the war in China and in Europe).

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

September, 1944
Volume 65, First Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under

the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

SEPTEMBER, 1944

(Manuscripts are printed in the order of final acceptance)

Dr. Hull's *Principles of Behavior* 3
By Robert Leeper

Factors influencing the efficacy of punishment and reward: The opportunity for immediate review, and special instructions regarding the expected rôle of punishment 53
By J. M. Stephens and J A. Baer

Congenital alexia and its relation to optic perception 67
By Paul Schilder

Test scores for clerical aptitude and interests before and after a year of schooling 89
By Samuel F. Klugman

(OVER)

Copyright 1944, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879

Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The motivational value of barriers for young children 97
By Irvin L. Child and Elizabeth Adelsheim

A survey of psychological, social, and environmental differences between advanced and retarded readers 113
By Joseph Jackson

Behavior efficiency as a determining factor in the social adjustment of mentally retarded young men 133
By Sidney W. Bijou

An approach to a more comprehensive analysis of mentally retarded predelinquent boys 147
By S. W. Bijou and B. R. McCandless

An experimental and critical study of the intelligence of Negro and White kindergarten children 161
By Fred Brown

SHORT ARTICLES AND NOTES 177
Apparatus for the study of delayed reaction in children (Robert O. Scow and Margaret E. Robins) 177

BOOKS 181

CRITICAL REVIEWS OF RECENT BOOKS 183
L. S. Hollingworth. *Children Above 180 IQ.* (Reviewed by W. Drayton Lewis) 183

BOOKS RECENTLY RECEIVED 193

The Journal of Genetic Psychology, 1944, **65**, 3-52.

DR. HULL'S *PRINCIPLES OF BEHAVIOR**

Department of Psychology, University of Oregon

ROBERT LEEPER

A.

If the principle of closure has any vitality, this book ought to go over with a bang. It is a book which has been awaited with unusual interest, not only by those who have been impressed by Clark Hull's earlier work, but also in the sense that it fulfills some clearcut needs for a certain type of book on the psychology of learning. To see this book in proper perspective, therefore, it may be well first to define the reasons for both this active and this potential interest in it.

American psychologists have tended to look on learning as more clearly a worthwhile field of research than any other division of general experimental psychology. One reason for this is the practicability of the field for research purposes, especially as contrasted with the relatively more intangible fields of motivation, thinking, and personality. But learning has appealed also because of its practical importance. For, in a real sense, learning is the chief concern of any parent, educator, psychotherapist, athletic coach, advertiser, or political leader. It is no wonder, then, that in American psychology, in almost all of our discussions of child psychology, social psychology, educational psychology, etc., we relate the rest of our interpretations to some general principles of learning. Some notable books have not followed this pattern—witness the recent excellent books on psychotherapy by Carl R. Rogers, Frederick Allen, and S. R. Slavson—but these exceptions are generally books that reflect clinical work rather than experimental work.

The widespread interest in problems of learning, therefore, is one reason why Hull's book will appeal. A second reason is that Hull's book represents the two lines of interpretation of learning that have been by all odds the most favored by American psychologists—the two relatively objectivistic theories of learning derived originally from Pavlov and Thorndike. In this respect it is interesting to contrast Hull's book with two others—John A. McGeoch's *Psychology of Human Learning* (Longmans, Green,

---

*This paper was begun as a book review, but became too important for just that function.—*Ed.*

1942), and Edward C. Tolman's *Purposive Behavior in Animals and Men* (Century, 1932). Both of these are remarkably capable books. But neither of them is adapted to American psychology as successfully as is Hull's book. McGeoch's book is respected because of its tremendous command of experimental material and because of its clear and reasonable interpretations. But, with its strong emphasis on rote memorizing as *the* great field of research on learning, it was published 30 years too late. It does not appeal to us as covering enough territory or as giving an adequate general theory of learning. Tolman's book, with its emphasis on trial and error learning and on animal experimentation as the means of clearing up many basic theoretical questions, is really more at home in current American psychology. But Tolman's interpretations have apparently not been convincing to a majority of American psychologists. I have just checked on six outstanding recent textbooks in general psychology—Boring, Langfeld, and Weld's *Introduction to Psychology,* L. E. Cole's *General Psychology,* J. F. Dashiell's *Fundamentals of General Psychology,* F. Ruch's *Psychology and Life,* Shaffer, Gilmer, and Schoen's *Psychology,* and R. S. Woodworth's *Psychology.* Tolman is not listed at all by three of these books, and is cited only in insignificant ways by two others. His interpretation of learning is mentioned with some adequacy only by Woodworth. Furthermore, none of these six books except Woodworth's makes any mention, in their discussions of learning, of several research workers whose experiments are highly consonant with much of Tolman's interpretation—namely, K. F. Muenzinger, I. Krechevsky, D. K. Adams, and K. Zener. Lashley's work on cerebral mechanisms is typically cited, but there is virtually no mention of his considerable body of experimentation on learning as a behavioral phenomenon.

If these six texts indicate the predominant viewpoint of American psychologists (as they probably do, judging both by the eminence of their authors and by their wide adoptions), they indicate that Tolman's work has not produced any sizable deflection of older trends of interpretation of the nature of learning. Just as N. R. F. Maier found that problem-solving tends at any given moment to have a "direction" that interferes with the use of any knowledge not consonant with that direction, so in American psychology our attempts to solve the problems of learning have been dominated (perhaps justifiably, of course) by the directions implicit in the theories of Pavlov and Thorndike. The influence of studies of learning by such workers as Tolman and Lashley has not passed a certain threshold value required to jolt our interpretations of learning out of their old orientation. Partly, therefore, the appeal of Hull's book will come from the fact that it conforms with the direction of interpretation which American psychologists now generally find it possible to take.

Not only does Hull's book build on the foundations of Thorndike's and Pavlov's interpretations, but it tries to clear these interpretations of any faults and tries to integrate them into one unified, systematically developed theory of learning. Most psychology texts, on the contrary, have given a double interpretation of learning, presenting on the one hand the concepts of conditioning and presenting on the other hand the concepts of learning derived from the trial and error point of view. This situation has not been satisfying, but it has not been evident how a single, unified interpretation could be developed. Some psychologists, of course, have avoided this duplication by virtually disregarding one tradition or the other. But this solution has not appealed to most psychologists. The notion that all learning can be reduced to conditioning principles has received some enthusiastic speculative application, but it has not distinguished itself as a means of formulating significant research problems for widely different types of learning situation. Thorndike might have captured the loyalty of most American psychologists with an opposite view that all learning is reducible to trial and error concepts. But, though ingenious, his recent experiments have been unconvincing as a basis for any general theory of learning. They have been impressive in bulk, but they have suffered from the great fault of too narrow a sampling of types of learning situation, types of subject, and means of measuring products of learning. Furthermore, Thorndike's theory has become increasingly vague in its effort to meet various criticisms. The "response," for instance, that is "connected" with a situation (29) can be anything—a pattern of muscular movement, a sensation, percept, memory, mental image, idea, judgment, etc. This vagueness protects the theory against some criticism, but at the cost of too much of its vitality.

When Hull tries to weld these two lines of interpretation into one homogeneous system, therefore, he is trying something which American psychologists long have wanted. And when he uses the earlier and more nearly objectivistic concepts of Thorndike, rather than some of his more recent concepts, Hull likewise is following a tradition which still is favored by a large portion of American psychologists.

However, to have a wide appeal, a book must of course not only conform to tradition in certain respects but move ahead of the procession in other respects. The chief point in which this is true of Hull's book is its stress on the idea (7) that psychological research should be ". . . directed by systematic and integrated theory rather than by isolated and vagrant hypotheses." Hull has won distinction for himself as one of that small but increasing number of psychologists who try to work by an explicitly formulated method of systematically developing a coherent theory, deducing particular experimental implications of that theory, and testing these implications experimentally. In contrast, too much of the work of the rest of us has been either a systematic gathering of data without any preliminary definition of problems to be tested, or a matter of experimental testing of isolated hypotheses—witness the jumble of topics under almost any name in the *Psychological Register*, for example. But we have not been proud of this procedure; and increasingly, as psychology has matured, we have come to grant the justice of Hull's demand for research that would derive its problems from a coherent theory rather than just from inspiration or from chance suggestions from some other experimenter's work.

Some psychologists, of course, have thought that Hull has been too interested in formulating methodological principles and too little interested in using them in experimental work. However, any full consideration of his work ought to take account of the experimental research of his students, because, just as in the case of Kurt Lewin, the bulk of his experimental work has come through the studies of his students rather than by his own hands. Actually, the experimental literature of the last decade has contained a large number of excellent studies of learning by Hull's students, with problems clearly defined, techniques carefully worked out, and with careful statistical treatment of results. Even by its summary of these particular studies the present book could well claim some value as a research monograph.

The appeal of the book will be not alone to research workers on learning, however, but also to workers in other fields who want an authoritative, definitive summary of the principles of learning applicable in other areas. It is for this group, at least in part, that Hull intended the book. Thus, the Preface starts as follows:

> As suggested by the title, this book attempts to present in an objective, systematic manner the primary, or fundamental, molar principles of behavior . . . the present work may be regarded as a general introduction to the theory of all the behavioral (social) sciences.

Indeed, two books have already been published from the Institute of Human

Relations at Yale that base their interpretations of learning almost entirely on Hull's concepts (N. E. Miller and John Dollard's *Social Learning and Imitation*, Yale Univ. Press, 1941; and J. W. Whiting's *Becoming a Kwoma*, Yale Univ. Press, 1941). So, apparently, Hull's theory has appealed convincingly to at least some of his colleagues as having the wide applications in social science that he feels it might have.

Many of the detailed features of the book, too, are likely to increase the interest in it. In contrast with the impossibly formidable *Mathematico-Deductive Theory of Rôte Learning* of 1940, the present volume is quite readable. The style suffers a bit from pompousness (it would rather say "when light waves impinge on the visual receptor" than "when light comes into the eye"), but otherwise a considerable clarity of exposition marks the book. A sizable amount of experimental material is cited, much of it summed up by graphs in which curves have been fitted mathematically to the empirical values. The closeness with which the curves fit the empirical data tends strongly to give a general impression of accuracy of interpretation. Especially technical parts of the discussion are segregated at the ends of the chapters. Excellent summaries conclude all the chapters, and the final chapter summarizes the whole book. These devices give an excellent means of getting a clear idea of the theoretical structure which Hull has developed.

All in all, therefore, probably the great majority of those who read the book will get a strongly favorable impression from it. It will serve, in all probability, considerably to lengthen the life of the approach to learning which it endorses. I will be very much surprised if we fail to find a great amount of quoting of this book, in the next 10 or 15 years, as the outstandingly authoritative and tough-minded treatment of the general theory of learning.*

B.

If Hull's *Principles of Behavior* lived up to the rosy expectations described above, a review of the book ought properly to be both brief and sweet. This review, however, will be neither. For, much as I sympathize with Hull's viewpoint in certain respects, and much as I admire some features of his method, I am aghast at the product of his work. It creates a good first impression. But when one spends on it, not the amount of time which most psychologists can afford to give to it, but enough time really to test its logic and to check its empirical material against the experimental literature, one flaw after another appears, and one is driven finally to the conclusion that most of the major conclusions of the book, experimentally speaking, are either meaningless or demonstrably unsound.

---

*See footnote, p. 52.

I admit that this statement must not sound reasonable. And, unquestionably, when the author of a book has the eminence and background of solid accomplishment of a man like Hull, the burden of proof is surely on the reviewer who expresses such an opinion. I will try to indicate the evidence, therefore, with considerable care.

I do not imply, however, that this book will not have made a contribution to the field of learning. Even if the book does have many faults, it has one main virtue which Hull wished it to have: it provides a coherent, logically developed system which facilitates the task of passing judgment on the broad approach to learning which Hull endorses. The really important question in the following review, therefore, is not the question whether Hull's book has certain faults or virtues, but is the question whether a "peripheralist" explanation of learning, such as tends to be derived from the fundamental concepts of Pavlov and Thorndike, gives us a good basic approach to this field. Admittedly, any theory of learning cannot be more than a rough approximation now, since our knowledge is so incomplete. But it is important for us to decide, and I think it is possible for us to decide, whether we are going in the right general direction when we adopt the "direction" used by Hull, or whether we can already see the need of a different approximate statement of the nature and conditions of learning. Hull's book can help us significantly in making this decision, because indirectly it can help us to see to what an extent the peripheralist view of learning has been a philosophy rather than a viewpoint sensitive to experimental evidence.

Hull, of course, has been one of the most vocal psychologists in demanding that we must be rigorously clear in our definitions, explicit in our logic, and humble in our treatment of experimental evidence. Thus, as descriptive of experimental spirit, one could hardly find a better statement than occurs in this book (p. 12):

> . . . empirical observations are regarded as primary, and wherever a generalization really conflicts with observation the generalization must always give way.

But, in the actual execution of the book, vagueness is often the rule; and, in his derivation of hypotheses from the experimental literature of learning, Hull is not guided mainly by experimental evidence, but by the philosophical faith that relatively mechanical concepts *must,* for some very general reasons, ultimately prove to be correct. So strong is Hull's faith in the fundamental soundness of his position that he preserves a stony silence even toward work by Lashley conducted explicitly to test some of the concepts Hull builds on.

This willingness to subordinate experimental evidence to deductive faith is of course no new thing in psychology. Nor is Hull the only psychologist, by any means, who accepts the particular deductive faith that relatively simple and mechanical processes must be the truth. In fact as I. Jenkins has said in a paper on the *Postulate of an impoverished reality* (13):

> . . . all distinctively modern thought, and so a large part of . . . contemporary theory and practice, is based upon a single postulate . . . the postulate that reality and nature are poorer, more bare, than our experience of them . . . that nature can be explained entirely in terms of matter and motion; that these alone are real. . . .

The reason why psychologists tend especially to sympathize with this sort of viewpoint is that, for a long, long time, the processes of the mind were conceived as mysterious processes that had no resemblance to the sequences of cause and effect which a science might investigate. This notion, one might say, was the bad Giant who lived in the land and who devoured all the embryonic psychologists. Then, along came the youthful scientific viewpoint and slew the Giant (except that he did not really die, so that it has been necessary to keep on slaying him every once in a while). The weapon which the young giant-killer used, so it is claimed, was the concept that psychological processes, far from being mysterious or inscrutable, are really only rather simple or mechanical connections between receptors and effectors. And, consequently, from Descartes on down, some thinkers have felt that the hope of truly scientific work in psychology depends most, not on following experimental evidence wherever it shall lead, but on sticking to this idea that all the seeming complexity of psychological processes is illusory and that, really, some fairly mechanical processes account for all the psychological facts that actually are facts. As time has passed, psychology in many respects has outgrown this simple faith. But, in 1943, in Hull's *Principles of Behavior,* the hands of the clock turn back, and Don Quixote rides again.

We could leave the book to establish itself, therefore, as a piece of charming fiction, with nice pictures, nice printing, and everything. But the trouble is that most of our general psychology texts tend to follow the same tradition which this book sets. If Hull has written fiction or philosophy, it also is true that he has numerous fellow-writers in a field which aspires to be a field of science, instead. It may be, though, that by exaggerating this tendency to rely upon a deductive faith rather than upon empirical evidence, Hull may again perform the service that Cervantes' hero rendered in another field.

C.

Before attempting to substantiate the criticisms suggested above, however, we need to get a general picture of Hull's interpretations.

We need not tarry over the first sixth of the book. The first two chapters present Hull's familiar (and really excellent) abstract methodological concepts, and the three subsequent chapters lead up to the discussion of learning by sketching such matters as adaptation by behavior, the receptor-connector-effector mechanisms, and innate behavior. A few stimulating slants are suggested. Thus, Hull emphasizes the idea that, for drives or needs to be adaptive, they must be able to arouse behavior of many different forms, with certain responses most likely to be evoked, other responses less likely to occur but still possible enough, etc. However, there is not much in these early chapters to justify a description of the book as other than a treatise on learning.

The main part of the book deals with two main problems. The one division (Chapters VI-XI, XIX, and part of XVI) deals with conditions governing learning or habit-formation. The second division (Chapters XII-XVIII, except part of XVI) deals with conditions governing the *use* of habits.

In both of these divisions, the theory which Hull presents is a *peripheralist* theory of learning. By this I do not mean merely that it is an objectivistic theory, built solely on data on stimuli and effector responses. It is such; but so likewise is the theory of learning of E. C. Tolman, and yet Tolman's theory is clearly a *centralist* interpretation.

What is meant by describing Hull's theory as peripheralistic is that, fundamentally, he interprets habits as relatively simple and mechanical connections between the two peripheral events of stimulation and effector reaction, and that, moreover, he believes that the conditions governing both the formation and the use of habits are relatively simple and mechanical conditions. To be sure, Hull almost certainly would deny this charge. He could point, for instance, to the fact that he speaks of the nervous activity involved in behavior as "this extraordinarily complex process" (p. 29); and he could point to the fact that his concepts of afferent interaction, reactive inhibition, summation, etc., obviously represent complications of the basic idea that the nervous system serves as a connecting link between receptors and effectors.

It is true, indeed, that the phenomena of learning which he describes are more complex than those which sufficed to satisfy John B. Watson. But, as the following summary will reveal, the various portions of his theory justify the description of his theory as portraying learning in relatively simple, mechanical terms.

1. *Hull's Delineation of the Nature of Habits*

At different points, Hull uses four different expressions for designating a habit, alternating between them as though they were approximately interchangeable. For the most part, the crucial terms of these four expressions are not defined. It is apparently assumed that the terms are sufficiently clear from ordinary usage. However, it cannot always be assumed that Hull wants his terms to be understood in their ordinary sense. For instance, in some of his formal postulates he uses the terms "afferent neural impulse" and "afferent impulse" (singular) in a sense which cannot possibly refer to the ordinary usage of this term (see Postulates 1 and 2, p. 47; Postulate 4, p. 178). Granting, then, that there is some uncertainty in what Hull intends because of this lack of definitions, it may be said that Hull uses the following four means of designating habits:

First, habits are spoken of most commonly as stimulus-response connections. Second, habits also commonly are spoken of as receptor-effector connections. Thus, after discussing reinforcement phenomena, he states:

> This central fact shows conclusively that reinforcement leaves within the organism a relatively permanent connection *between the receptor and the effector associated in the original reinforcement. It is this which in the present system is meant by the term "habit," a technical adaptation of* the common-sense concept that goes by the same name (p. 117, italics added; see also pp. 69, 315).

This seems like a definitive statement. But, unless a stimulus is the same stimulus only when applied to the same receptor, and unless a response is the same response only when executed by the same effector organ, this designation of habits as "receptor-effector connections" is not the same as their designation as "stimulus-response connections." Nor is the concept the same as that implied in the following statements from the start of the chapter (p. 183) on "stimulus generalization":

> . . . every reinforcement mediates connections between a very great

> number of receptor and effector processes in addition to those involved in the reinforcement process. . . .
>
> 1. The reaction involved in the original conditioning becomes connected with a considerable zone of stimuli other than, but adjacent to, the stimulus conventionally involved in the original conditioning; this is called *stimulus generalization.*
>
> 2. The stimulus involved in the original conditioning becomes connected with a considerable zone of reactions other than, but related to, the reaction conventionally involved in the original reinforcement; this may be called *response generalization.*
>
> 3. Stimuli not involved in the original reinforcement . . . become connected with reactions not involved in the original reinforcement . . .; this may be called *stimulus-response generalization.*

Really, this statement should be taken as Hull's chief description of the products of learning, because it points most definitely to the experimental facts which he would defend as best established.

In the fourth place, however, Hull speaks of habits as connections between afferent and efferent neural impulses (p. 404), afferent and efferent neural discharges (p. 111), or an afferent impulse (singular!) and a reaction (p. 178). In fact, he refers to this as the most satisfactory characterization of all, on the ground that the central afferent discharge is determined not only by the particular stimulus, but also by "afferent neural interaction" from other afferent impulses. However, Hull does not implement this concept with any principles that would throw light on the question of how greatly an afferent discharge might thus be modified and still evoke the related reaction, or on the question, either, of what factors give rise to much afferent neural interaction. Indeed, Hull says, ". . . under most circumstances there is a close approximation to a one-to-one correspondence, parallelism, or constancy between" the stimulus and the central afferent discharge on the one hand, and between the efferent discharge and the effector reaction on the other, so that these pairs of terms ". . . for purposes of coarse molar analysis . . . may be used interchangeably" (p. 112). Nor is this concept of afferent neural interaction presented as explaining the phenomenon of stimulus generalization (p. 354). So, practically speaking, although Hull presents this as his most technically-adequate designation of habits, the concept is not made ready for use.

All of these designations of habits, however, are somewhat comparable in that they are all peripheralistic in their emphasis. When an animal learns, it does not acquire "knowledge," "expectations," "hypotheses" or any other such thing, but acquires connections between events at the periphery.

The only departure from this is provided by the concept of afferent neural interaction, which is a move in the direction of a Gestalt interpretation of nervous functions. But since this latter concept is not developed in a way which permits any experimental predictions (other than the negative warning that you may not get the results which you should expect from the other principles!), it leaves Hull's designation of habits a peripheralistic view.

2. *Hull's Description of the Factors Governing Habit-Formation*

The same tendency to describe the phenomena of learning in relatively simple and mechanical terms is seen in Hull's principles regarding the forming of habits. According to Hull, "connections" are established merely because an afferent discharge arrives in the central nervous system at about the same time (but not necessarily prior to the time) that an efferent discharge occurs, and because all this "is closely associated" with a reinforcement or need-reduction (p. 178). Nothing is said about the way the learner perceives or organizes the situation, or about any processes of mental set, thought, or sensory re-organization. In fact, very little is said about any sort of contribution by the learner other than that he serves as a sounding board, as it were, for the various stimulus-determined processes.

Hull carries out this theory with a consistency which makes his meaning unmistakable. He says, for example, that in a conditioning experiment connections are formed between the response and *all* of the stimuli acting on the learner's receptors. Thus, Pavlov's dogs must have formed habits connecting the salivary response, not solely with the reputed conditioned stimulus, but also with postural stimuli, visual stimuli from the room, olfactory stimuli from the restraining harness, bites from the dog's fleas, etc. (pp. 205-206). The connections between most of these stimuli and the response are likely to be weak, however, because (*a*) many of these stimuli are relatively weak, and (*b*) their time of onset is generally too remote from the time at which the response and reinforcement occur. Furthermore, these incidental connections are soon subjected to extinctive inhibition. The fleas unwisely bite the dog on many occasions when no food is given; the odor of the harness continues between reinforcements. Hence by relatively mechanical processes the conditioned response tends eventually to be connected effectively only with the stimuli to which, under ordinary life conditions, it ought to be connected to be biologically adaptive. However, *if these corrective conditions were not present, all stimuli in the situation would be connected effectively with the response.*

Although Hull suggests no reason why this might be expected, he devotes one chapter to the proposition that the first occurrence of a learning situation makes the greatest contribution to the development of the habit-strength attainable under the given learning conditions, the second occurrence of the learning situation produces an increment of habit strength equal to the same fraction of the potential habit strength still not achieved, etc. This is what Hull means by saying that habit strength is a "positive growth function" of the number of trials. This is not intended as meaning, however, that the habit will express itself in any observable way in the early trials, because Hull assumes a certain inertia or reaction-threshold and also assumes certain chance variations ("oscillation" factors) which would operate solely to *obstruct* the occurrence of the response (note this odd assumption, pp. 326-328, 393). In view of these additional factors, Hull argues that an *S*-shaped performance curve may be secured during learning even though the habit strength is growing maximally at first (pp. 326-335).

Several other factors also are related quantitatively to the rate of learning. Thus, this rate is determined partly by the amount and nature of the reinforcing agent. The increment of habit strength from any given trial is greater the smaller the interval between response and need-reduction. In fact, Hull says, if the interval between the response and reinforcement is greater than 30 seconds there is not likely to be any learning except in situations in which connections have been built up between "fractional antedating goal reactions" and some of the stimuli in the learning situation. When an animal has been conditioned to secondary reinforcing agents, however, these operate to produce a learning of whatever stimulus-response connections they are closely associated with (pp. 89-98).

The rate of learning also is influenced by the time relations between the conditioned and the unconditioned stimuli. The most favorable time relation is that which exists when the signal precedes the adequate stimulus by "something less than a half second" (p. 176). As the time relations vary from this optimal point, in either direction, the rate of learning falls off sharply. In fact, says Hull (p. 177), if the stimuli do not set in operation some major physiological cycle,

> It is doubtful if true trace conditioned reflexes can be set up when the onset of the unconditioned stimulus follows the termination of the conditioned stimulus by more than about three seconds.

Only one experiment is cited for this conclusion, and in this one experiment the curve anterior to this optimal point, even including measurements for backward conditioning, is as high as the curve posterior to it.

Referring back to these several factors affecting habit-formation, Hull (pp. 388-389) says in the final chapter of the book:

> There are doubtless other conditions which influence the magnitude of the increment of habit strength resulting from each reinforcement. Those listed above certainly are typical and probably comprise the more important of them. An adequate statement of the primary law or laws of learning would accordingly take the form of an equation in which ${}_SH_R$ would be expressed as a joint function not only of N but of the quantity and quality of the reinforcing agent, and of the temporal relationships of S to R and of R to G.

If anyone wishes to get from Hull's theory of learning, then, a means of planning efficient learning situations, or if anyone wishes to find an explanation of why learning occurs easily in some situations and only with great difficulty in other situations, these are the only factors to which Hull's discussion would direct his attention.

The development of extinctive inhibitions is really another form of learning, however, and mention of it should be linked with the above discussion. Hull's theory is that not all of experimental extinction is due to extinctive inhibition (or, conditioned inhibitions). Instead, from every response, whether reinforced or not, there is produced a certain "need for rest" or "need for cessation of that activity." This "reactive inhibition" spontaneously dissipates within a few hours, rapidly at first, and then more and more slowly. However, if trials are massed, it accumulates to significant amounts, and when each response ceases there occurs a satisfying or reduction of this need to cease responding. Hence, a habit of not responding, or of conditioned inhibition, is learned. Therefore, learning not to respond does not depend on any perception that the situation has been changed or on any intelligent observation or hypothesizing that the signal is no longer followed by its former referent; but the learning not to respond depends merely on the same simple conditions of S, R, and need-reduction that produced the original learning. Spontaneous recovery from experimental extinction may occur insofar as extinction is a product of "reactive inhibition," since the latter (p. 298)

> ". . . dissipates spontaneously through the passage of time, but ${}_SI_R$" (the conditioned inhibition), "being a true habit, presumably does not, at least to any great extent."

3. *Hull's Description of the Factors Governing the Use of Habits*

Really the largest portion of the book is devoted to this problem of the factors governing the evocation of habits. In addition to the two factors just mentioned (temporary states of reactive inhibition and enduring habits of extinctive inhibition), the following factors are stressed:

(*a*). According to the principle of *stimulus-generalization gradients,* the power of a stimulus to evoke an habitual response decreases "according to a negative growth curve" as the stimulus differs more and more from the exact stimulus involved in the learning. With intensity differences this loss is much less than with qualitative differences (differences of pitch, e.g.).

(*b*). Special effects occur from combining different stimuli. Because of external inhibition, an alien stimulus presented with a conditioned stimulus will lessen the effect of that conditioned stimulus. When several stimuli are presented together, each of which is positively connected to the same response, summation will occur, except that (p. 219):

> If two stimuli separately and equally conditioned to the same reaction are later presented simultaneously, and their afferent neural interaction effects are sufficiently great, they will jointly evoke a weaker reaction tendency than would be mediated by either one of the stimuli acting alone.

No criteria are given for knowing when the afferent interaction effects will be "sufficiently great." This means, I cannot refrain from inserting, that compound conditioned stimuli may either summate or interfere, with the conditions altogether unspecified as to which of these will occur. It is natural, therefore, as Hull says, that some findings by Shepard and Fogelsanger, cited by Kohler, ". . . present no paradox whatever" (p. 219) for his principle of compound stimuli!

(*c*). An active motive or drive must be present for a stimulus to evoke its related habit. Evocation is a multiplicative, not an additive, function of drive and of habit strength. However, this drive need not be the one present in the original learning; any drive tends to activate all habits through hypothesized chemical or hormone-like factors that operate directly on the nervous system. These substances also act specifically on particular parts of the body (e.g., nutritional factors on stomach muscle), and hence they produce drive stimuli which tend to be conditioned to particular responses.

(*d*). According to the principle of *behavioral oscillation,* the reaction caused by a habit will vary considerably even under conditions that seem constant. This oscillation is caused by fluctuations of the stimulation thresholds and latencies of individual neurones and muscle fibers, and also by spontaneous discharges of impulses from neural centers. Such oscillations are presumably uncorrelated in different parts of the reactive mechanism and might exert a joint influence which would appreciably change the character or the intensity of a response. This "indeterminateness" phenomenon in psychology, as one might call it, is particularly the reason why a massing of statistical evidence is ordinarily required in behavioral research.

On all of these matters regarding both the formation and evocation of habits, Hull has attempted to determine the quantitative nature of the mathematical relationships involved. In a few cases he concludes that the relationships are probably linear. More generally, however, his formulations are those of a "positive growth function" or its converse, a "negative growth function."

It might seem as though, with learning and habit-use dependent on such relatively mechanical considerations, the products of learning would generally be maladaptive, or at least unadaptive. However, Hull takes pains to point out the biological adaptiveness of the phenomena he hypothesizes. Repeatedly he insists that, for behavior to be efficiently adaptive, it must display such features as he describes. Thus, discussing stimulus generalization, he says (p. 389),

> *This fact . . . is of immense adaptive significance; since stimuli are rarely if ever exactly repeated, habits could scarcely function adaptively without it.*

Such statements are expressive of the fact that the book seeks to deal with learning, not just in terms of laboratory experimentation, but as something fundamental for real life adjustments.

D.

In the sections of the review following this one, we will be compelled to face some rather serious faults in Hull's book. But the book also has some of the fine qualities which we would have expected in it. Indeed, it is the outstanding worth of the book in some respects that makes the faults of the book the more regrettable, because in some ways it has the makings of a truly great book on learning. I will not dwell at length on these virtues, because they are readily apparent to any reader. But they should at least be listed.

First of all, Hull should be given credit for clearly defining a number of truly worthwhile problems of learning. It is important to ask how the number of reinforcements affects the strength of a developing habit, how habit strength is affected by the nature and amount of the reinforcement, and how learning is affected by temporal relations between signal, response, and reinforcement. It is important to explain experimental extinction. It is important to ask how the overt use of habits depends on stimulus generalization, conditions of motivation, etc. These are all good problems, and I believe that all of them are more clearly defined now because of Hull's book.

Hull, secondly, has presented some good experimental material on these problems; and even back of this, he also deserves great credit for having suggested and directed a considerable portion of the experimental studies which are reported at most length in the book.

In the third place, Hull has worked out the inter-relations between a number of concepts in a way which long has needed to be done. Thus, as mentioned above, he has faced explicitly the problem of why only the intended signal, rather than all of the stimuli in the conditioning situation, becomes an effective conditioned stimulus. His discussion of the rôle of motivation in conditioning has helped to bring conditioning and trial and error learning together. He has not merely echoed the common treatment of learning in terms of a mixture of principles from different sources, but has unified the principles into one homogeneous system.

Fourth, Hull has performed one of the main services which he insists that a systematist should serve: he has stated many of the concepts sharply enough, and has designated experimental implications of those concepts sharply enough, that he greatly facilitates the task of judging empirically the validity of the peripheralistic point of view in learning. In consequence, even if he has not been willing himself to consider the evidence which often disproves those implications and thereby refutes his major principles, he nevertheless has formulated the theory in a form which compels the theory to face the evidence. So, the book does demonstrate in some ways the usefulness of Hull's methodological concepts even though his personal devotion to some concepts prevents him from following their guidance throughout.

E.

In the present section, the discussion will present some concrete examples of the various types of fault found in the book. Important though these faults may be, however, I do not feel that the discussion should be left at that. Sometimes it happens that, though an author is vague in many of his definitions, careless in his use of experimental evidence, and more or less illogical, he still achieves a theory that contains so much good that one cannot afford to neglect it. With Hull's book, however, I am driven to the conclusion that virtually the whole theoretical structure is unsound. If this is true, it means that a number of our most common interpretations of learning are also unsound. Accordingly, in the next section, I will consider Hull's main principles one after another, and give the reader the opportunity to draw his own conclusions.

As a background for the discussion of the various types of fault in Hull's book, let us quote Hull himself:

> The essential characteristics of a sound scientific theoretical system. . . .
>
> 1. The *definitions* and *postulates* of a scientific system should be stated in a clear and unambiguous manner, they should be consistent with one another, and they should be of such a nature that they permit rigorous deductions (7, p. 495).
>
> 2. The labor of deducing the potential implications of the postulates of the system should be performed with meticulous care and exhibited, preferably step by step and in full detail . . . (7, p. 495).
>
> 3. The *statements in the theorems* should agree in detail with the observationally known facts of the discipline under consideration (8, p. 5).

Odd as it may seem that Hull should frequently violate these precepts after formulating them so clearly and emphasizing them so strongly, these violations are the main faults of the book. I will speak specifically of these points, except that I shall make no separate comment on the matter of whether he has been sufficiently logical. After all, the best proof of the logical pudding is the eating.

Before discussing the first and third criteria of a scientific system, however, let us add a fourth which is so simple that Hull, of course, took it for granted: the author of a theoretical system, when he reports experimental evidence, should report it in such a way that he will give a reasonably careful reader a correct impression of the experimental facts.

The several failures of Hull to achieve this fourth requirement of scientific work are not nearly as important as are his omissions of experiments which contradict his principles. But they are straws that show the way the wind is blowing; and since they are relatively simple points we might begin with them.

As a first example, let us take the presentation of two experiments by Hovland which constitute the sole evidence offered in support of the principle of stimulus-generalization gradients. Figure 42 (p. 185) sums up the data from one of these experiments. In this graph, "amplitude of galvanic skin reaction" is plotted against just noticeable differences from the tone used in the learning. The center of the base line is marked *"O,"* and then in each direction from this point the base line is labelled up to 200 *j.n.d.'s* of pitch. The high point of the curve falls above the zero point, and from this high point the curve falls away with an impressively accurate fit and with wonderful symmetry in both directions. It is the type of material in Hull's book which especially tends to impress a reader. As one psychologist has already expressed the matter to me, "What makes me think the book is important is the success with which Hull has formulated quantitative relationships. After all, his curves *do fit!*"

I happened to look up Hovland's article, however, to check on another matter. What was my surprise to find that Hovland's article showed no such double-winged curve! Half of his subjects (total $N = 20$) had been conditioned to the highest of four tones; the other half to the lowest of the four. Then, to avoid any artifacts from absolute qualities of the tones, all of the subjects had been thrown together, averaging all scores to the "conditioned tone," regardless of whether this was the highest or the lowest tone, then averaging all scores to the tone 25 *j.n.d.*'s away from that conditioned tone, etc. This procedure is legitimate enough. But it yields only a single curve, and what Hull did was to take this one curve and draw it in both directions. Not only that, but the title of Hull's graph states: "Note that the gradient extends in both directions on the stimulus continuum (vibration rate) from the point originally conditioned." Furthermore, as one of the four conclusions from the data, Hull states (p. 185):

> The pooled results from twenty subjects are shown in Figure 42. There it may be seen that: . . .
>
> 2. This diminishing generalization tendency extends symmetrically in both directions. . . .

One can only remark mildly that a mirror-image usually is symmetrical with its source. The original data *could* have been plotted to show the falling-off in both directions. But the fit would not have been so nice.

The same sort of treatment is given the data from Hovland's other experiment in Figure 43 (p. 186). Still later, in discussing the "stimulus generalization of extinction effects," Hovland's data on extinction phenomena are presented in this same unwarranted mirror-image fashion (Figure 58, p. 264).

As a second example of poor presentation of material, we may look at the sole bit of concrete material cited from the extensive data on learning provided by K. S. Lashley. After discussing some of Pavlov's observations on external inhibition, Hull (p. 218) says:

> What appears to be the same thing but in an even clearer form has been reported by Lashley (5, pp. 140-141). Rats were trained to jump to a black card bearing a white triangle and not to jump to a similar card bearing a white circle of the same area. (Incidentally, Hull does not even cite this correctly, as the negative design was a plus sign.) After the animals had attained more than 95 per cent of correct choices, new cards were substituted which contained the same figures, with four small figures added. The score of "correct" choices with the new cards fell to 90 per cent.

No measure of reliability of the difference is given, but Hull continues, using the conventionally modest style at the risk of implying that the interpretation is favored by Lashley:

> The interpretation is that the four additional figures so changed the afferent impulses . . . that the effective habit strength at the command of each was appreciably weakened. . . .

Since I knew that Lashley often has found it instructive to do rapid exploratory work with small groups of subjects, I looked up the reference. I found that this fall from "more than 95 per cent" to 90 per cent was based on 20 trials for each of two rats—just enough to warrant Hull's use of the plural! And, Lashley speaks of the item as indicative of "successful transfer to the other figures." Now, has Hull misrepresented or not? If all this had occurred in a senior thesis, we might say "No, he has not become sufficiently familiar with statistical safeguards." But when a man of Hull's known statistical skill proceeds in such a fashion, what can one say?

Another example was apparent to me because it concerned a study of my own which closely paralleled a slightly earlier study by Hull. In both experiments, rats were trained with simple two-route mazes in which they needed to traverse one arm when hungry and the other arm when thirsty. The motivation was alternated, daily in Hull's experiment, somewhat more irregularly in my own. The only vital difference of method in the two studies seems to have been this: In Hull's maze, when a rat traversed the wrong arm, its run led to a door which blocked access to the single end-box and the rat was compelled to retrace and go through the other arm to reach the end-box. In my own experiment, on the other hand, since it was planned under the stimulation of Tolman's and Lashley's theories of learning, the maze had been constructed with two end-boxes, one containing food and the other a pan of water. When a rat made an incorrect choice in this maze, it entered an end-box containing the goal-material *not* desired on that day; and there it was confined for about a half-minute before its next run. The situation on correct runs was approximately the same, naturally, in both mazes. Therefore, *provided such confinement in the wrong end-box was not rewarding,* both experiments, according to Hull's theory, should have yielded similar learning curves. On the other hand, according to such a theory as Tolman's, the two experiments should have yielded extremely different results because, on the wrong trials, the rats should have acquired useful knowledge when they were confined in the end-box containing the goal-material not desired on that day and, in Hull's maze, they should have acquired some knowledge that would confuse them when they were required on the following day to select the path that had led to the empty blind alley.

The results of the two experiments were these: For the first month of training, Hull's rats averaged less than 50 per cent correct on the first runs of each day. They improved gradually, but required 25 periods of training of eight days each before attaining 80 per cent correct on first runs. In my own experiment, however, only *one* such 8-day period was required before reaching 83 per cent correct on first runs in the second 8-day period; and 94 per cent was attained in the third 8-day period. To account for this difference Hull says in his book:

> . . . thirsty rats supposedly satiated with food will, *after receiving even a few drops of water,* very generally eat if food is available (6, p. 270). . . . Thus after the first trial Leeper's animals were presumably operating under both drives, and one drive or the other was reinforced no matter which path was traversed (p. 235, italics added).

The reference cited is his own related article. But the only factual statement on the matter contained in it is the following statement, which almost surely refers to the effects that occur when a rat is returned to the living cage and is permitted to drink to temporary satiation (9, p. 270):

> . . . very thirsty animals seem not to eat dry food in normal amounts. As a rule, at once after securing water rats will proceed to the food jars and eat greedily even though the same jar of food may have been standing in the cage some hours prior to the giving of water.

This quotation therefore hardly gave him a basis for saying what rats would do in the maze after receiving only a few sips of water. And, actually, on the incorrect runs the rats almost never touched the undesired goal-material, but spent most of their time vigorously trying to push up the lid and escape from the end-box. I regret that my article failed to specify this fact. But, considerable evidence was mentioned which shows indirectly that the rats were not "operating under both drives" after the first trial. Thus, in the days 6-26 of the training, during which the per cent correct on first runs was 86 per cent, the per cents correct on the 3rd, 4th, and 5th runs for the same period were 86, 91, and 85. A curious reversal effect appeared on the 2nd trials, but it seems to have other explanations. Furthermore, when the rats later were given 20 trials in immediate succession, the hunger drive held up rather evenly throughout, and the thirst drive fell off only as indicated by the following percentages of correct runs in successive groups of five trials: 100, 97, 80, 53.

A control experiment was performed, using Hull's type of maze, and Hull's results were duplicated for the period tested. And on the other hand, using an elevated maze with three end-boxes, and adding some "latent learning" conditions, I had other rats reach 100 per cent scores on first trials within five days on each of three problems with such hunger-thirst alternations. So, it seems quite safe to say that it was exceedingly hampering to the rats in Hull's maze to meet a door blocking access to the one end-box, and that it was definitely helpful to the rats, in the other mazes, to come into an end-box and find there the goal-material not wanted at the time. This in turn suggests that learning is produced at least partly, not by reinforcement, but merely by an opportunity to perceive "what response leads to what consequences." Even though a crucial theoretical point is at issue, however, Hull did not cite the available indirect evidence that might have ruled out his explanation, nor did he ever write to inquire: "I am inclined to interpret your results as due to such and such a factor—does this assumption fit your observations of the facts?" He preferred to infer the empirical facts, justifying the inference by a citation which does not really say what he cites it as saying.

Even though there are, therefore, at least some cases of poor presentation of experimental material, this is not the most serious fault of the book. The most serious problems are the book's dependence on a small number of studies (often experiments with small numbers of subjects) and Hull's refusal to discuss, or even mention, a considerable number of expertly-conducted experiments which run flatly contrary to his main conclusions. Even an inspection of the index of authors' names is sufficient to reveal the one-sidedness of the evidence which Hull employs. Thus, to name some workers whom one might most expect in a book like this, there is no mention of the work of K. F. Muenzinger, W. J. Brogden, E. Culler, I. Krechevsky, H. Harlow, D. K. Adams, W. S. Hunter, R. Loucks, N. R. F. Maier, T. C. Schneirla, H. S. Liddell, A. Melton, J. A. McGeoch, or L. E. Cole. None of the experimental work of Tolman is cited, nor any of his interpretations of learning—only the ideas of "molar" concepts and of "intervening variables." Aside from one joint study, none of the experimental work of E. Hilgard or D. Marquis is mentioned. One slight reference is made to G. Razran's work. Two slight references are made to Lashley's work, including the one citation already discussed. No work of the Gestalt psychologists is mentioned except in a few vague, slurring, and often unfair references such as the following (pp. 379-380):

> Indeed, if current configurationism is ever formulated as a true scientific theory . . .
>
> It is hoped that the derivation of the major phenomena of stimulus-pattern learning . . . will contribute to the dissipation of current misunderstandings among psychologists, since these are a source of such deep and painful confusion to the scientific public. However, optimism in this connection is seriously dampened by the conviction that the differences involved arise largely from a conflict of cultures . . . which, unfortunately, are extra-scientific and are not ordinarily resolvable by either logical or empirical procedures.

By actual rough count of the space devoted to the direct discussion of experimental methods and results of work on learning, my estimate is that about 75 per cent of such space is devoted to material from Pavlov's laboratory and from Yale (about 60 per cent from Yale). Only about 25 per cent goes to all of the rest of the experimental literature on learning! Sometimes, however, not even studies from Yale are cited. Thus, Chapter XIV depends heavily on studies by two of Hull's students—S. B. Williams and C. T. Perin. But no mention is made of another study by another of Hull's students, conducted between the times of the other two, and seemingly quite as trustworthy as either of them—the study by R. S. Sackett (24). It would not have supported the elaborately-analyzed findings derived from the other two studies, but I can detect no fault in it except that it tends to spoil the nice picture yielded by the other two experiments. Similarly, although for years Hull employed the study by himself and M. J. Bass (10) as proof of gradients of stimulus generalization, his present chapter on stimulus generalization makes no mention of it, and makes no attempt to reconcile its opposite findings with the data from Hovland.

No more need be said now, however, regarding such omissions. Since the experimental literature on learning is so huge, an author naturally cannot cite everything. When we discuss the various principles of Hull's system, however, we will see that the chief charge is the omission of directly relevant material. Before concluding the present section, a third fault of Hull's book may be illustrated—its lack of explicit definition of crucial terms.

Some features of the book might give the appearance that Hull has defined with care. Thus, the book ends with a glossary of more than 100 technical symbols used in it. But this glossary merely gives the verbal equivalents of these symbols and does not, in turn, define these verbal equivalents. Nor in the text itself are the crucial terms sufficiently defined. For example, as has been mentioned above, Hull apparently intends that learning should be understood primarily as something that produces connections of a sort that yield the phenomena of stimulus generalization, response generalization, and stimulus-response generalization. If this is so, the definition of these terms is vital. But, for instance, all that we are given in the definition of response generalization is the statement:

> The stimulus involved in the original conditioning becomes connected with a considerable zone of reactions other than, but *related to,* the reaction conventionally involved in the original reinforcement . . . (p. 183; italics added).

Now, what does "related to" mean? Must the same muscles be used for a second response to be "related to" the original response? Is a response "related to" the original one if it produces the same effect on the environment? Or, if "related to" means neither of these things, what is the criterion of relatedness?

Again, Hull says that the best designation of a habit is that it is a connection between a response and an afferent discharge as modified by afferent interaction. If such is the case, it is desirable that the principle of afferent interaction should be operationally meaningful. But this postulate has this formal statement (p. 47):

> "Postulate 2. All afferent neural impulses active in the nervous system at any given instant, interact with each other in such a way as to change each" (each neural impulse—note the singular) "into something partially different in a manner which varies with every concurrent associated afferent impulse or combination of such impulses."

This is an assertion, in other words, that there is a phenomenon of afferent neural interaction. But, what are its conditions? What is the difference produced in "each" afferent neural impulse? How much are afferent impulses modified, and what factors produce much or little effect? Unless such things are stated, the principle is almost meaningless, of course. But the only answer given in the remainder of the Postulate is this:

> Other things equal, the magnitude of the interaction effect of one afferent impulse upon a second is an increasing monotonic function of the magnitude of the first.

This sentence illustrates Hull's usual flair for some sort of quantitative statement; but it likewise says almost nothing. It is easy for us to believe that there are significant afferent interactions: the findings with binocular rivalry, the phi phenomenon, reversals of ambiguous figures, distortions in visual illusions, etc., are abundant testimony that there must be some such effect, even though Hull characteristically mentions none of these. But to explain such afferent interactions, Hull has probably seized on one of the most minor factors of all, but then asserts (p. 46):

> This hypothesis makes possible the explanation of many important behavior phenomena, otherwise inexplicable, such as . . .

One wonders what the word "explain" means to Hull! Actually, if Hull *does* attempt to develop this concept in a way that will provide the basis for definite predictions, he will find himself treading many paths already traversed by Gestalt psychologists. Their principles of perceptual organization need to be stated more explicitly, it is quite true; but even now there is some real guidance in their principles of factors governing figure-and-ground organization and other types of grouping; and on the other hand it would be hard to find in psychology a principle that is more nearly meaningless than this one of Hull's.

If, therefore, it is hard either to prove or disprove Hull's theory at certain points, part of the difficulty is that sometimes it cannot be said what that theory is.

F.

In the present section we get down to the fundamental question: *"Is Hull's theory of learning a good theory?"* This question is not necessarily answered by the materials given above. I believe they put Hull's theory under suspicion, but they are not fatal to it. We need now to make a frontal attack on the problem. We need to go down the line examining one after another of Hull's main principles of learning.

1. *Hull's Principles Regarding the Nature of Habits*

As we have seen, Hull has used four different means of designating habits. Thus, habits are receptor-effector connections: ". . . reinforcement leaves within the organism a relatively permanent connection between the receptor and the effector associated in the original reinforcement. It is this which in the present system is meant by the term 'habit' . . ." (p. 117). But even a beginning student in psychology should be able to spot the inadequacy of this as a designation of habit. The experiments are legion that show that habits still function despite changes of locus of stimulation and of muscles involved in the response.

Secondly, when Hull sums up his series of chapters on habit-acquisition, in his Postulate 4, he states that whenever an "effector activity" and a "receptor activity" are closely associated in time with each other and with either a primary or secondary reinforcement, ". . . there will result an increment to a tendency . . . for that afferent impulse on later occasions to evoke that reaction" (p. 178). As brought out earlier, this sort of expression is vague because Hull does not define the crucial terms "effector activity" and "reaction." Hence we cannot be sure what different reactions would still be the "same reaction" and would tend to occur as a consequence of that habit. However, let us take the loosest meaning which the term "reaction" might have in a peripheralistic theory, and let us see whether, even with this meaning, Postulate 4 may be reconciled with experimental data. That loosest meaning would assert that reactions are the same, regardless of whether they involve the same muscles or the same pattern of movement, provided they produce similar effects on the environment.

To determine whether Postulate 4 can be accepted, even in this loose sense, let us consider the numerous expertly conducted latent learning experiments (27, 4, 5, 2), none of which Hull mentions. These have been criticized, of course, on the grounds that the supposedly unrewarded runs actually are rewarded through a satisfying of some exploratory drive. Suppose for the moment we accept this criticism as valid—what, then, do these experiments show? The fact remains that, until some appropriate goal-material has been found at the end of the maze, the "reactions" of the rats are reactions of wandering into blind alleys, leisurely washing themselves, sniffing at cracks, etc. If Postulate 4 is correct in asserting that reinforcement establishes connections between receptor activities and the reactions shown in the learning situation, and if it is true that these supposedly unrewarded runs have been rewarding, the conclusion follows that the effect of that learning should have been the forming of connections so that later the stimuli from the maze and from the hunger condition should have evoked reactions equivalent to those wandering, fur-washing, crack-sniffing reactions. The minimum conclusion, however, that may be drawn from these latent-learning experiments is that what is learned is not connections that tend to produce *effector activities* or *reactions* similar to those shown in the learning period itself, but is something that may produce more or less opposite reactions.

As stated previously, however, the designation of habits which seems most favored by Hull is neither of the above, but is the concept that ". . . every reinforcement mediates connections between a very great number of receptor and effector processes in addition to those involved in the reinforcement process . . ." (p. 183).

Hull has reason to say that there are a "very great number" of such connections formed, if this theory is correct. For example, when a person learns to read some new words, he forms connections (on the stimulus side) of such a nature that he now would know how to pronounce and interpret the words as presented in many different styles of type, many different sizes of type, and even many different colors. The number of such connections, on the stimulus side, would presumably equal the number of discriminations which the person could be trained to make regarding such differences of style, size, color, and illumination of the printed words because, according to any peripheralistic theory, such differences would have to be differences of stimuli, or else discrimination habits could never be established.

On the response side, too, the number of connections would be enormous. The experiments on equivalence of response illustrate this. For example, Heinrich Klüver reports that, after a Cebus monkey had learned to pull in food by means of one object, it readily transferred to the use of a sack to whip in the food, a piece of rope for the same purpose, a cardboard triangle, a piece of wire netting, and a chain. When only a newspaper was available, and with that pinned down by a heavy box, the monkey tore off pieces of the paper, worked and rolled them into a partly rigid mass, and used this (14, 15).

Not only this, but according to Hull's theory it also would be necessary to say that connections must be established between the stimuli and patterns of movement which the person has not yet learned to make. Thus, when a person learns to spell a word as a child, he is forming connections between the sound of the word and reactions of typewriting, telegraphy, deaf-mute finger language and whatever other patterns of movement he may acquire years later!

Now, admittedly, it is possible to say that a learning situation causes the establishment of thousands or tens of thousands of connections and cross connections, even to patterns of movement not yet acquired by the person. But one of the qualities of good scientific theory is economy, provided the more economical formulation can account for the facts. Thus, Copernicus's helio-centric theory of the movements of the planets was based on the same data employed by Ptolemaic explanations in terms of cycles and epicycles with reference to the earth as a fixed point of reference; and it was only the quality of economy, and hence of greater intellectual efficiency, that led astronomers to renounce the Ptolemaic views for those of Copernicus.

The alternative to Hull's multi-connection theory of the nature of habits can account for the same facts by four processes, the first described by the principle that, as E. R. Guthrie (3, p. 18) has expressed it, "Somewhere in the brain the impulses from receptor elements can be combined to act as functional units." That is, the process in the central nervous system begins with a process of organization of nervous impulses. The factors governing this organization are not simple. Partly the organization is determined by the pattern of stimulation of the receptors and by the innate properties of the nervous system, as the Gestaltists have emphasized most strongly; but partly the organization is determined also by habits of sensory organization, by motives active at the time, by mental set, by temporary conditions of fatigue such as Köhler has hypothesized to account for fluctuations of reversible figures, etc. Admittedly, such processes of organization are constructs, known only by inference; but this would be no objection from Hull's point of view, because he sees clearly that psychology must deal in terms of inferred processes. So, as the first portion of an alternative theory, there need to be various principles that would enable us to know what sorts of organizations are established that would serve as the functional units (just as a chemical compound is a functional unit) that would provide the basis for the further steps in the process of response.

The second aspect of an alternative theory would be that, after a habit is established, such a pattern of afferent activity tends to arouse, redintegratively, the same further pattern of brain activity (or, even more widely, of physiological activity) which previously was evoked only by the more adequate conditions present in the learning situation. Thus, the same motivational process tends to be aroused that was aroused in the learning situation, the subject tends to visualize the same situation that he perceived earlier, etc. However, third, rather than these redintegrations appearing in a pure form, another step usually occurs of fusion or integration of those redintegrative processes with the perception of the current situation. Then, in the fourth place, the redintegrative process as thus modified is used by the subject as the basis of whatever effector activities are at his command and seem appropriate to him as a means of dealing with the particular objective situations.

According to such a formulation, habits might be described particularly as connections between patterns of stimulation and central redintegrative mechanisms. Or, we might use either Tolman's (30) expression, that a habit is a "sign-gestalt-expectation," or Ralph White's (31) simpler language, that habits are knowledge of what leads to what. Let us admit here, too, that these are gross terms, molar terms. But again that need not disturb Hull, for he undertook to describe in this book ". . . the basic molar behavioral laws . . . the uniformities discoverable among the grossly observable phenomena of behavior as contrasted with the laws of the behavior of the ultimate 'molecules' upon which this behavior depends . . ." (p. 17).

In such an alternative theory there is no imputation of peculiarly human characteristics to animals. The redintegrative mechanisms are always more or less restricted in their adequacy of representation of what effects will be secured by what techniques. Thus, when Klüver's (14, p. 297) Cebus monkey was baffled by a problem of food suspended out of its reach, it directed its behavior towards the floor, and "went through the motions of raking in food with a *T*-stick although there was no food lying on the floor, or she did the same with a ring, a sack, or a wire." Too, as time passes, such redintegrative mechanisms tend to become less elaborate and less detailed than they were originally. They become more nearly simple perceptual-motor connections. But such phenomena do not controvert a description of habits as "knowledge" or as "redintegrative mechanisms," but merely define some of the fruitful problems of learning which must be well explored under this approach.

Naturally, one assumes a somewhat complicated process when he assumes that animal behavior depends upon a fourfold process of perceptual organization, redintegration, fusion of redintegrative material with current perceptual material, and final effector activity. But animal behavior is in many respects marvelously flexible and adaptable; so we are driven to recognize either a series of such processes or tens of thousands of cross connections such as Hull conceives to be established by each reinforcement.

2. Hull's Principles of Habit-Formation

We may sum up in nine main points the principles of Hull regarding the acquiring of habits. If this theory is sound, these nine principles should harmonize with available experimental evidence.

(1). Hull asserts that *all adequate stimuli favored by proper temporal relations become connected with the response,* if associated also with proper reinforcement. What is learned is not something determined by highly variable perceptual organizations within the learner, but is primarily stimulus-determined.

It is not surprising that Hull cites no experimental evidence to support this principle. It would be an endless task to prove that all stimuli in a conditioning situation become connected with the response. This principle is a key feature of Postulate 4, therefore, not because of empirical evidence, but because his general approach demands the idea deductively.

In contrast with this absence of empirical support for the principle, however, we may easily find experimental evidence that denies it. Thus, in an experiment reported by Lashley in 1942, four rats were trained to jump to the *larger* of two circles, this learning requiring an average of 95 trials. They were then tested with the smaller circle paired with a larger triangle. They always chose the larger figure (the triangle). Then 200 further trials were given with the small circle still negative and the larger figure, the triangle, positive. According to Hull's theory, the rats should certainly have learned a connection between the triangular form and the positive jumping response, because (*a*) the triangle certainly stimulated the rat's eyes, (*b*) rats can learn to distinguish triangles from circles, and (*c*) the triangle as a stimulus was presented only with proper time relations to the response and reinforcement.

However, when the rats were next presented with a triangle and a circle of equal areas (the same triangle, and a circle intermediate between the original two), the rats responded merely with persistent position preferences. Retrained with the large triangle and small circle (with no errors in 20 trials), the rats were then tested with a small triangle and large circle. All rats chose the large circle rather than the triangle in 20 out of 20 critical trials. As Lashley (19, p. 259) comments:

> "If the animals are given a set to react to one aspect of a stimulus

> situation" (as, here, to relative largeness), "large amounts of training do not establish association with other aspects, so long as the original set remains effective for reaching the food."

Another striking disproof of this first principle of Hull is an experiment reported by I. Krechevsky (17).

The evidence points, therefore, not to the mechanical sort of interpretation offered by Hull, but to an interpretation which Lashley has expressed in this way (19, p. 242):

> 1. The mechanism of nervous integration is such that when any complex of stimuli arouses nervous activity, that activity is immediately organized and certain elements or components become dominant for reaction while others become ineffective. This constitutes a "set" to react to certain elements. Such an organization is in part described by Gestalt principles of perception, in part by principles of attention.
>
> 2. In any trial of a training series, only those components of the stimulating situation which are dominant in the organization are associated. Other stimuli which excite the receptors are not associated because the animal is not set to react to them.

(2). Hull asserts that *the ordinal position of each trial determines in a relatively simple way the increment of habit strength from that trial, and that, more specifically, habit development is a positive growth function of the number of trials.*

This is a sweeping conclusion, a conclusion of the sort that requires extensive evidence before it may be accepted. But Hull presents only three studies that support this conclusion of Chapter VIII. An entomologist would not be content to present three species as evidence of the proposition that all arthropods have six legs; but Hull expects us to respect his book as scientific when he presents three specimens for a generalization covering such a large and heterogeneous field as the field of learning.

It is quite probable that there are many learning situations in which the greatest amount of learning occurs on the first trial, less on the next, and so on. This may well be true especially of those situations which are fairly well structured perceptually or cognitively at the very beginning (as in memory experiments with simple visual forms, lists of words, or stories). In order for a principle to be scientifically useful, however, we must know the conditions under which it holds and the conditions under which it no longer applies. To discover such limits of application, a theorist is under obligation to search for likely exceptions and to report them carefully.

There are such exceptions to Hull's principle. For example, opposite results were reported by I. Krechevsky (18) in a study designed explicitly to test this view of Hull and of K. W. Spence. Three groups of rats were trained with the jumping apparatus. The control group received consistent training throughout—Card *A* was always positive, and Card *B* always negative. With a second group, however, Card *A* was negative and Card *B* positive during the first 20 trials, and only after these first 20 trials were the cards used as with the control group. Still another group had its first 40 trials in this reversed fashion before ending its training with the cards arranged as for the control group.

According to his principle of the importance of hypothesis-processes in learning, even by animals, Krechevsky predicted that the rats given the reversed training would not show a greater total learning time, even including these reversed trials, provided the original period was not too long. According to the theory of Hull and Spence, however, these first 20 (or 40) trials should have developed connections between the "wrong" designs and the jumping and avoiding responses, and after the significance of the cards was reversed, the rats should have to tear down some old habits and build up the new connections. The reversed groups should therefore have required a larger number of trials to learn, *even excluding the trials before reversal.* This would be even more true of Hull's present theory. When Krechevsky performed this experiment, Hull and Spence (25, p. 431) were hypothesizing that the true learning curve is probably *S*-shaped, with ". . . the relative strengthening effect of a single reinforcement . . . least at the beginning and end of the learning and greatest in the middle portion." But now, since Hull asserts that the greatest increment to habit strength comes from the first conjunction of stimulus-response-reinforcement, the next greatest increment from the second such conjunction, etc., it would be even more true that Hull's theory would predict great differences of learning time between the control group and experimental groups.

Krechevsky found that the control group required 162 trials to learn; the first experimental group required only 141 trials, including the 20 trials with reversed training; the second experimental group required 171 trials, including the 40 reversed trials! Krechevsky concluded that, in about the first 30 trials, the rats must not have been learning to respond to the stimulus patterns as such.

Spence (26) has severely criticized these results. First, he asserted, it is not valid to condemn the Spence-Hull theory because the experimental groups took fewer trials after the reversal than the control group required, since the training technique tended, especially in the early trials, to eliminate position preferences. This objection is a refusal to face squarely the implications of their theory, because Spence and Hull hypothesize that proper conjunctions of stimulus, response, and reward cause increments of strength of the *S-R* connection from the beginning of training. Position preferences might mask the operation of that *S-R* connection at first, but would not obstruct its development. To discard the first 20 (or 40) trials for the control group would be utterly inconsistent with the logic of Spence's and Hull's theory. Second, Spence (26, p. 280) asserted,

> . . . Group III did require more trials than the control group . . . a result which would appear to be opposed to the non-continuity theory. . . .

But, asserting this, Spence made no mention of the size of the difference (162 as against 171 trials!), nor of the fact that the difference was only .63 times its standard error. Just as Hull's book abstains from giving any standard errors, so Spence spoke merely of "The fact that Group III learned more slowly . . ." (p. 281). Third, Spence said, the rats probably jumped without having the differential stimuli within their field of vision. But the rat's field of vision is quite large, the rats surely looked at the cards, and correct reactions could have been based on any small portion of the type of cards used, as Lashley's work has shown. Fourth, Spence (26, p. 281) said,

> . . . the relevant cue aspects simply were not received at first by the subjects. They did not "see" the rows of black squares because they were fixating other aspects of the stimulus complex.

Now, if this statement means "received" and "see" in the sense of "central afferent receptor discharge," as per Hull's theory of learning, the rats surely did "see" the stimuli. But, if by "see" Spence meant something like "pay attention to" or "have become dominant in perception," this fourth point is not a criticism, but is merely a concession of the whole Lashley-Krechevsky concept.

Now, we have no knowledge of whether Hull refused to mention Krechevsky's experiment because of any feeling that Spence had disposed of it. If such was the case, Hull depended on a grossly inadequate critical paper. On the other hand, if he realized the unsoundness of Spence's criticisms and yet still refused to mention this evidence against his own conclusions, it is another significant example of Hull's unwillingness to introduce his readers to any discordant evidence.

What happens on any given trial, therefore, is a much more complex matter than Hull suggests. Thus, even in relatively simple discrimination experiments, the early trials may produce no knowledge at all that the "crucial" stimuli indicate how the reward may be attained; instead, these early trials may contribute to the final mastery of the problem only by giving a necessary period in which misleading "hypotheses" may be tested out and eliminated.

(3). Hull asserts that *reinforcement or need reduction is essential for learning.* This law of reinforcement is, of course, the same as Thorndike's recent formulation of his Law of Effect. And this principle, either in its earlier or later forms, is almost universally accepted in American psychology. So, at least here it might seem as though Hull is on safe ground.

This concept is a ticklish one to discuss. For, in some important ways, rewards do influence learning and performance. When a hungry rat finds food at the end of the maze and eats the food, this event tends to establish a knowledge or expectation of the consequence which may be attained by traversing the maze. Such goal-expectations, in turn, tend to produce a more energetic exploration on subsequent trials and thus supplement the simple motive of hunger as a means of producing the exploratory activity indispensable in trial and error learning. Provided that what we mean by "situation" is the combination of external stimuli and of hunger, we surely can assert therefore, as does Hull's Law of Reinforcement, that a reward tends to cause a change in the learner of such a sort that, when the same situation returns, the rewarded responses will be more likely to occur.

However, Thorndike's Law of Effect and Hull's Law of Reinforcement are more than the assertion that, when some conduct has been followed by a reward, the learner is more likely to engage in the same conduct when placed back in the same situation with the same motivation as before. That would be a fairly limited and guarded statement. But these principles go beyond that. They assert, first, that learning does not occur except as rewards or need-reductions occur. They assert, second, that what is learned as a result of the rewards is not such central effects as goal-expectations or knowledge of what actions lead to what consequences, but is a connection between the situation and the *responses* that led to that reward.

To answer the question of whether such interpretations are correct, we should consider first the experiments on latent learning. These have been the object of some criticism, of course, so that most psychologists have not considered them disproof of the Law of Effect. Thus, Ruch (23, p. 295) says that the question of the latent-learning experiments ". . . is whether there was learning without drive"; and since there was presumably an exploratory drive, he concludes that the "principle of satisfaction" is supported, that "In the absence of such satisfaction we usually find little or no learning" (23, p. 294). But, although the Law of Effect implies that motives are essential to learning, it is not such a law of motivation. When we point out that motives are needed to produce a persistent, vigorous attack we are not thereby proving that *satisfaction* of those motives is necessary for learning.

Or, still more commonly, the critics of the latent-learning experiments have argued that these experiments do not disprove a Law of Effect because, during the supposedly unrewarded trials there is not only a motive of drive to explore, but there is also a satisfaction of this motive. This criticism thus takes one or two logical steps, and then stops. What happens if we go farther? Suppose we ask such a critic to explain the fact that, with a food-reward given from the beginning, the errors in a maze decrease. "Simple," he would say, "the errors fall because the rat is learning to run the maze correctly, and he is learning to run the maze correctly because, when he goes down the true pathway, he is soon thereafter rewarded."

"All right," we counter, "now explain this case: if rats never get food in the end-box, their error curves do not fall."

"Simple enough again. The errors do not diminish because the rats did not learn, and the rats did not learn because there was no more reward associated with going down the true pathway than there was with going into blind alleys."

"That's interesting. Now, these latent-learning experiments. Why does the error curve fall so precipitously after, say, the tenth trial, on which the rats first encountered food in the end-box? Doesn't this seem to indicate that learning really must have been occurring in the preceding trials, even though the conditions were not present there for evocation of that habit in a particular form?"

"Yes, but you forget that even though there was no reward of hunger there was presumably a rewarding of some exploratory drive."

"Where was that rewarded?"

"That's a foolish question. Everywhere, of course—when the rat went into the end-box, down the true path, in the blind alleys—everywhere."

And with that we might leave that conversation, because it illustrates the common logic fully enough. To wit: satisfaction of hunger causes a learning of responses that particularly led to that satisfaction, but satisfaction of the exploratory drive leads to a learning, not of the responses that led to that satisfaction, but of responses of traveling the maze with few errors!

However, the assertion may be made, still further, that rewards are not essential to learning. As bearing on this, I might refer to the pair of experiments by Hull and myself that have been described earlier, on the hunger-thirst alternation. In these experiments the mazes were so simple, and so many trials were given, that it hardly seems appropriate to say that any "satisfying of an exploratory drive" could have been of much relative importance after the earliest days of training. In both mazes, it will be remembered, when the rat made a correct run it had practically the same experience of coming into an end-box containing whichever goal-material (food or water) was needed on that day. So, any differences of results cannot be explained by effects of rewards. But on the wrong runs, the rats in Hull's maze came to a door which blocked their access to the single end-box, and in the other maze they came into a second end-box which contained the goal-material not desired then. These incorrect runs must have produced the difference of results in the two experiments. In both experiments, despite Hull's inference to the contrary as regards my experiment, the wrong runs were not rewarded, but were punishing or frustrating. Nevertheless, Hull's rats required 25 8-day periods of training before attaining a criterion of 80 per cent correct on first trials; my rats required one 8-day period. It must be, therefore, that when the rats saw the goal-material not desired at the time, even though that situation was frustrating to them, they learned something that helped them to make an opposite selection of paths on the following day when their motivation was shifted. And the rewards on correct runs must not have stamped in the responses that led to that reward, but stamped in some sort of habit which did not interfere seriously with such shifting of response when the motivation was changed. So, it may be assumed, rewards cause learning and tend to cause a recurrence of the rewarded response on subsequent trials only because of a more inclusive phenomenon that "the process of responding to a situation tends to establish a knowledge of the properties of the situation."

To make such a distinction, and on the basis of it to object to the Law of Reinforcement as inadequate, may seem like splitting hairs. But science often proceeds best by splitting hairs—by making distinctions that look pointless to a commonsense view. So it does not matter that the Law of Effect (or Reinforcement) is not "all wrong." It is wrong in the sense that a law is wrong in physics that asserts that "Heavier materials fall faster than lighter materials." My own faith is that psychology is now sufficiently mature that it cannot remain content with such grosser generalizations, but ought to make the finer distinctions that clearly are possible.

(4). Hull asserts that *secondary reinforcing agents* (stimuli previously closely associated with primary need-reductions) *cause a learning of stimulus-response connections with which they are conjoined in time.* Thus, he says, the elimination of the early blind alleys in a maze cannot be accounted for by the reinforcement from the food in the end-box, because too much time intervenes between the responses to those alleys and the final reinforcement. Instead, the correct responses to the early parts of the maze are acquired because various stimuli are present early in the maze which have in turn been associated with the primary reinforcement. The walls and ceiling, e.g., are like the walls and ceiling of the end-box. Hence these serve as secondary reinforcing agents themselves, and also call out various "antedating goal reactions" (such as chewing movements, perhaps) which also provide some secondary reinforcement.

This mechanical explanation sounds plausible until we ask a few questions. For example, when are these secondary reinforcing agents present? The answer must be: "All through the run." Then we ask: "If such is the case, the secondary reinforcement must be associated with retracing and with entrances into blind alleys just as truly as with movement down each segment of the true path—why, then, does the secondary reinforcement help the rat to learn to take the correct path?"

After all, as stimuli, the blind alleys are more similar to the end-box than is the true pathway. Of course, an entrance into a blind alley is not followed by any primary need reduction. But neither is the passage down a segment of the true path! Intuitively, of course, Hull knows that the two experiences have different kinds of effects on the rat. But he cannot derive the difference from his principle of secondary reinforcement, and Hull's system is supposed to be logically explicit, not an affair of intuitive predictions.

If, on the other hand, we hypothesize that rewards are not the necessary condition of learning, but that learning comes fundamentally from a superposition or close succession of several perceptual processes, regardless of whether any of these have a rewarding quality, there are no special difficulties in the question of how learning occurs, say, at the start of a maze.

(5). Hull asserts that *punishment produces learning only through the need reduction which occurs when the punishment ends.* This is one of the most clearly doctrinaire concepts of Hull. In many learning situations the pain does not end until long after the response which produced the pain. When a person scalds his face, for example, the pain does not end soon. Or still more clearly, if a parent by carelessness permits a small child to be scalded, it is long after the accident before there would be any reduction of the punishment or discomfort aroused in the parent. Such a great temporal separation of need reduction from the response ought, according to Hull's scheme, to prevent learning. Still, this would be a powerful learning situation. A punishment produces learning because it causes one clear, vivid perception of certain potentialities or properties of the situation. If that perception is clearly superimposed on another clear perception of what action or type of situation led to that punishment, the conditions of learning are met regardless of when the final need-reduction occurs.

(6). Hull asserts that *the important aspect of a reward is the "need reduction" which it brings, rather than any positive quality of the reward experience.* On this concept, too, it is necessary to speak carefully. It is of course true that rewards are "need reducing." And, as E. R. Guthrie especially has emphasized, this need reduction is what brings the goal-seeking or exploratory activity to an end. Furthermore, when I say that the positive value of a reward is the chief reason for the influence of rewards on learning, it may seem as though I am contradicting what was urged above, that rewards are not essential for learning.

There is a big gap in Hull's theory in this connection. Why should a "need reduction" produce learning? On this question Hull makes no suggestions at all. It favors his relatively mechanical approach to say that what is crucial about a reward is some decreasing of impulses from some drive stimulus, rather than some positive value of the reward; but he has no evidence for stressing this one phase of the correlated event rather than the other.

When we look afresh at empirical material, the evidence seems to point to the positive influence as crucial, however. For example, when you offer a piece of candy to a small child and let him take a tiny taste of it, you are reducing the need for candy in the sense that, if he took perhaps a thousand times as much candy as he got in his hasty lick, he would be content to stop eating candy. But in every other respect you have *aroused* a need rather than reduced it—the child may have been busy with his toys the moment before, with no active motive of craving candy, but now the craving for candy has been excited as an active process and his behavior becomes dominated by the knowledge that you have some candy and that you are withholding it from him. The positive effects of a reward are important in conduct in such a case, not any need reduction. The old Roman banquets may violate our aesthetic sense, but they were based on a correct formulation that behavior is directed, not at need reduction, but at satisfaction in a positive sense, with need-reduction, as they saw the matter, only an unfortunate after effect.

(7). Hull asserts that *a relatively simple formulation may be given regarding the influence on learning of temporal relations of the conditioned and unconditioned stimuli.* Thus, citing only one experimental study of trace conditioned responses, Hull concludes that the most favorable time relations exist when the signal precedes the adequate stimulus by somewhat less than half a second, and that, as the relative temporal placement of the two varies in either direction from this optimal point (even if this creates the conditions for "backward conditioning"), the gradient of influence on learning is about the same. Indeed, he says (p. 177):

> It is doubtful if true trace conditioned reflexes can be set up when the onset of the unconditioned stimulus follows the termination of the conditioned stimulus by more than about three seconds.

Hull's treatment of this problem affords a beautiful example of how to develop simple conclusions by selecting evidence you want. Thus, to begin with, Hull presents the data only of Helen M. Wolfle's 1930 experiment, which yielded almost symmetrical gradients in both directions. He does not present the data from a later experiment by the same investigator. In that later experiment, for instance, none of the subjects showed conditioning when the sound preceded the adequate stimulus (a shock to induce finger withdrawal) by two seconds; but, in contrast, there was a considerable indication of apparent backward conditioning. Nor did Hull mention, secondly, that this very peculiar result makes it probable that her experimental conditions allowed considerable effects from self-instructions, which G. Razran (22) has shown are so important in experiments on conditioning with human subjects. In this connection, we must remember particularly that Hull does not conceive of a conditioning experiment as some special and peculiar example of learning—the question in such a study is merely, for example, "Can a person learn to lift his finger off a grill after a given signal if the signal precedes the shock by two seconds or more?" I submit that it is inconceivable that Mrs. Wolfle's subjects did not learn that the sound signalled the shock. It is inconceivable that they would not have lifted their fingers from the grill if it had not been for their efforts to avoid finger withdrawal unless it occurred "involuntarily" or in response to the shock.

In the third place, Hull declined to mention an experiment by A. L. Bernstein (1) in 1934, after Wolfle's studies. Hilgard and Marquis's book has cited the experiment prominently, so it cannot be that Hull overlooked the article. The experiment was apparently excellently performed, under L. E. Cole's direction. It used more subjects than Wolfle's at each interval studied, and employed also a more careful means of measuring conditioned responses. But Bernstein found almost no conditioning when the signal followed the shock; and with forward conditioning he found roughly equal frequencies of conditioned reaction for each interval from .25 sec. up to the longest interval employed (1.48 sec.). Bernstein concluded specifically that the data obtained from the longer intervals do not support Wolfle's hypothesis that the amount of conditioning in human subjects is a function of the time interval. Hull has therefore selected evidence on the basis of its agreement with his hypotheses, rather than on the basis of technical expertness of available studies.

Hull's whole treatment of perseverative stimulus traces is foggy and indefinite. As the basis for "after effects of receptor stimulation" such as he mentions (p. 41), there are at least five possibilities: (*a*) after-discharges from reflex centers that continue after the incoming impulses have ceased, (*b*) after-effects that are maintained by muscular activity and resulting proprioceptive stimulation, as in some of W. S. Hunter's delayed reaction experiments, (*c*) a thalamic-cerebral circulation of nervous impulses such as Cannon's theory of emotion suggests, (*d*) an intra-cerebral circulation of impulses, such as Lashley regards as both possible and probable, and (*e*) processes maintained by verbal activity of the learner.

Now, if Hull has attempted to develop merely a molar theory of behavior, he was not under obligation to explore these alternatives and to decide between them. But, if that is the case, it is equally true that he should not then have ruled out any of these alternatives and proceeded as though the after-effects of a stimulation can be carried only by some rather mechanical factors which trail off quickly in some physiologically-determined way. Actually, probably all five of the alternative means listed above should be recognized as being important for habit-formation, and rather than there being some mechanical rule of how after-effects diminish with time, what is more probable is that some after-effects disappear almost immediately and some may endure for hours or days. Thus, an expert workman in a factory will have very short after-effects from all of the normal noises in the factory, even when their physical intensity is great, but he will have long-enduring after-effects when he hears some much weaker noises which indicate that his machines are not operating properly.

(8). Hull asserts that *conditioned inhibitions are developed because the ending of each response reduces a need for ceasing to respond.* This concept is fairly important for Hull's theory. Extinctive inhibition is a matter of learning and habit; and if the learning of conditioned inhibitions cannot be explained on the basis of Hull's simple principles of learning, this failure would call into question his general explanation of learning.

Deductively, Hull does bring the development of conditioned inhibitions into his general scheme. Each response, he asserts, whether it is reinforced or not, produces a degree of "need for rest," or "reactive inhibition," which tends to dissipate fairly soon, but which acts like any other motivational state while it lasts. The ending of each response brings a reduction of this need. Hence, even during reinforced trials, the subject learns not to respond. This habit of not responding is masked under ordinary conditions of learning, because the habit of responding becomes strengthened at a faster rate than the habit of not responding. But during extinction training only the habit of not responding is developed, and this habit finally becomes the stronger of the two.

To support this interpretation, Hull presents only one experiment (pp. 292-293). He speaks of it, too, as failing to conform in some ways with the theoretical expectations. In contrast with this shaky evidence supporting his interpretation, various clear-cut experiments might have been cited to disprove it. One such is an experiment by J. S. Calvin which he cites as bearing on other points of the same chapter (pp. 290, 296). Calvin showed that, when reinforcements were given at the rate of three per minute, the frequency of conditioned lid reflexes rose to about 50 per cent in 50 trials, whereas with reinforcements given at the rate of 18 per minute the frequency hardly went above 20 per cent in the same number of trials. This is good support for Hull's idea of states of reactive inhibition. But, if the 18 reinforcements per minute developed a greater need for rest, they also should have developed indirectly a stronger habit of conditioned inhibition. But, after a 24-hour rest, which Hull elsewhere characterizes as sufficient to dispose of the reactive inhibition but as having little influence on conditioned inhibitions (p. 286), the frequency of conditioned responses rose to 71 per cent for the group given three reinforcements per minute, and to 72.5 per cent for the group given 18 reinforcements per minute.

A second approach would be this: Since this state of reactive inhibition dissipates within a few hours, conditioned inhibitions should develop only from massed extinction trials. Thus, if maze trials are given only once a day, the rats should hardly show any unlearning of their former responses after the reward is removed. Especially is this the case since Hull asserts that the secondary reinforcing agents, which would still be present, are quite important in such situations (p. 302). However, with such daily trials, the errors increase markedly within two trials after the reward is eliminated or is changed to a less-desired goal-material (30, ch. 3).

Third, Hull might have cited two experiments by Lloyd Humphreys, one of which was done at Yale under his own direction. In these two studies Humphreys found that, when only half of the trials had been reinforced during the training, with reinforcement scattered a bit irregularly, the two groups apparently became equally well conditioned in the same total number of trials (which is hard to reconcile with Hull's principle of the influence of number of reinforcements on learning), and, what is even more important, the two groups differed greatly in their rates of experimental extinction. After only several trials the group given invariable reinforcement showed a very rapid falling off of response. They had apparently formed an expectation of reinforcement on every trial, and when the reinforcements no longer came it was easy for them to form the new expectation that the situation had been changed. But the subjects given irregular reinforcement showed no falling off for four trials in one experiment, three in the other, and then only a slow decline thereafter (11, 12). Hull mentions this work by Humphreys only in another chapter, and only in the following precise fashion:

> It has been shown by Humphreys (7), for example, that the course of extinction is *rather different* when the reinforcements of the original learning have been accompanied by a considerable number of non-reinforced stimulations (p. 337; italics added).

(9). As a summary test of Hull's principles of habit-formation, we need to examine the view implied by Hull that *one can plan efficient learning situations if he takes into account solely the factors affecting learning discussed in this book—time relations between stimulus, response, and reinforcement, and the amount and nature of the reinforcing agent.*

In his narrow conception of the factors governing rate of learning, Hull of course reflects previous work on conditioning. Even in Hilgard and Marquis's splendid book on *Conditioning and Learning* one may search in vain for a discussion of factors determining rate of learning, speed of conditioning, etc. The problem is vital, nevertheless. Our principles have only minor theoretical or practical significance if we know only that learning can occur under certain conditions, but know very little about the conditions under which much or little learning will occur.

A good test of Hull's theory is given by J. R. Liggett's (20) experiment on olfactory discriminations. He used an apparatus in which the rats on each trial went down a tunnel through which strongly-scented air was pulled by a fan. If one odor was present, the rat had to turn to the right at the end of the tunnel to reach food, with another odor it had to turn to the left. The conditions of learning specified by Hull were fulfilled. There was a stimulus which certainly caused a central afferent discharge, there was a response and a closely subsequent reinforcement. Yet, with about a thousand trials, the rats did not develop any stable discrimination habits, and Liggett concluded that rats must have, at the most, but little sense of smell. In contrast, when H. G. Swann (28) modelled a smell apparatus on the principle of Lashley's jumping apparatus, he succeeded in establishing stable discrimination habits in rats within a few score trials. Such contrasts could be multiplied from many other studies.

Actually, there is sufficient empirical knowledge to justify a much more elaborate and definite set of principles of factors governing learning than Hull has given. Many of these principles cannot be stated in physiological terms. But this should not have deterred Hull, since he set out to develop a behavior theory that would be useful for the social sciences, and since he undertook to do this by developing "molar" principles, "uniformities discoverable among the grossly observable phenomena of behavior" (p. 17). For example, we need some "capacity laws." Thus, we need principles that relate differences of learning capacity to differences of species, age, and *IQ*. We need laws of apperceptive background, of motivation, of the perceptual organization of the learning situation, of the learner's mode of attack, and of mode of presentation of the learning situation. Such principles could lead to a really rich understanding of learning situations. They could not be derived from Hull's several principles of factors productive of increments of habit-strength; so they are just as primary as are his principles. Perhaps none of them are principles that can now be stated in terms of quantitative formulas. But, as has been seen, Hull has not actually been successful in this even with the few meager factors he has discussed, but has achieved the appearance of Science by the selection of "helpful" experiments. These other principles need not imply any lack of reverence for mathematics or quantitative treatment. But they do imply that the proper method of growth of psychological theory is a method of "approximation and correction," and not a process of refusing to consider anything which cannot already be expressed by some formula.

3. *Hull's Principles of Conditions Governing Habit Use*

The situation with regard to Hull's principles of habit-evocation is somewhat more satisfactory than with regard to the above-mentioned concepts of the nature of habits and of the conditions governing formation of habits. In the present matter, there are a couple points to praise, as well as some other points that need adverse criticism.

(1). *Hull makes a valuable contribution by his emphasis on the idea that separate problems are involved in the evocation, as contrasted with the acquisition, of habits.* To appreciate the progress which this represents, one need only compare his discussion with J. A. McGeoch's (21, pp. 597-599) treatment of the question of whether learning can be distinguished from performance. Logically, McGeoch says, the distinction may be made, but it is ". . . practically and operationally of little importance," since "The only way we can know that learning has occurred is by an observation of successive performances, since *learning is a relation between successive performances.*"

Now, different conditions call out quite different performances following the same previous learning situation. Therefore McGeoch would have had to say that any learning situation establishes innumerable "learnings." This concept is just as awkward as is Hull's notion that each learning situation establishes innumerable "connections" between variants of the original stimulus and variants of the original response. But on this matter Hull has moved on to the sort of approach which is more characteristic of contemporary science—the approach which defines the work of science as the organizing of knowledge, not just in terms of relationships between directly observable data, but partly in terms of constructs or inferred variables. We have always been willing to use some such inferences in psychology. In his clear-cut differentiation between problems of habit-formation and problems of habit-evocation, however, Hull has carried our common viewpoint into a more mature and complete form than has been common in our discussions of learning. His acceptance of this distinction should facilitate greatly the relating of his concepts to those of Lashley and Tolman, both of whom stressed this same distinction many years earlier.

(2). Hull asserts *that S-shaped curves during learning are due to masking of habit-increments by oscillation factors and by reaction threshold.* Hull had to reconcile his theory of maximum learning on early trials with the empirical observation that, in some learning situations, no signs of such learning occur until after a number of repetitions of the situation.

The solution is plausible as related to some learning, such as learning to give the names of objects, because in such cases any chance factors (oscillation factors) have almost no likelihood of helping to produce a correct response. However, *S*-shaped curves are also gotten in some discrimination experiments, as with Lashley's jumping apparatus. Hull does not consider this sort of situation in this connection. This is unfortunate, for it disproves his explanation as a sufficient interpretation, and would have helped him to search farther. For, in the first place, any threshold or inertia which might tend to obstruct a response has been overcome through the preliminary training in which the rat learns to jump to one card or the other. In the second place, since the oscillation factors would be as likely on any trial to favor the correct response as the wrong response, it would follow that, if the first trials produce some increments to habit strength, these increments should start to register their effects in performance immediately, in some appreciable improvement over the chance 50 per cent correct.

(3). Hull's concept *that grddients of stimulus-generalization may be accounted for by simple stimulus differences* seems to be an over-simplification. It must be true, of course, that there are limits to the extent of stimulus generalization; and presumably there must be some gradations of effect within these limits. But the probability is that such gradients are often products of relatively complex processes which can easily be overlooked when, as J. S. Bruner and Gordon Allport[1] expressed it, we prefer to settle questions of the nature of learning, etc., by use of animal subjects or "human subjects who for the duration of the experiment are totally mute."

This matter was first called to my attention by an unpublished study by Treva Jane Ogan under R. H. Waters at the University of Arkansas. She was repeating an experiment by K. S. Yum (32) in which it appeared that, in paired associates learning, when some stimulus words were replaced by less and less closely related synonyms, the subjects were less and less able to recall the associated response words. When Miss Ogan started to test her subjects for retention, she soon came rushing into Dr. Waters' office. "What am I to say to them?" she exclaimed. "They want to know whether they are supposed to give the word that went with the original word, or whether they should say nothing since the word is changed." "Well," replied Waters, "let's see what Yum did." So they ransacked Yum's article. He gave no hint of such a response from his subjects. So Waters told her, "Go back and tell them again, as Yum presumably did, that you are showing them the stimuli involved in their earlier learning, and that they are to give the appropriate associated words." Soon she was back again. "It doesn't work. They know every word that has been changed, and they insist on telling me that they are changed." It was finally decided to give them no help, but merely to confine the instructions to the ambiguous statement that they were to give the appropriate response words. Accordingly, some subjects interpreted their task as that of giving the word only when the stimulus was unaltered; some tried to give the response word to any stimulus recognizably similar to the original stimulus, and others assumed that they should give the paired word when the resemblance was close, but not otherwise. So, like Yum and like Hovland, Miss Ogan found a nice stimulus-generalization gradient. As far as can be told from rough inspection of Hovland's original scores for different individuals, it seems as though the same phenomenon occurred with his subjects—not in the sense that they "decided" to make galvanic skin reactions to some tones and not to others, but at least in the sense that different subjects formed different expectations of which tones were likely to be followed by shock. Hull's treatment of stimulus-generalization gradients should not have neglected such relatively more complex possible sources of this effect.

---

[1] Fifty years of change in American psychology. *Psychol. Bull.*, 1940, **37**, 766.

(4). Hull's principle *that motivation is essential for all use of habits* may be a fairly sound generalization with regard to skeletal-muscle responses of a gross overt sort. But it is a poor general principle. Actually, every object that we look at calls out redintegrative processes so that, as R. S. Woodworth expresses it, our "situation set" is constantly changing as the objective stimulation changes. This must surely be included as an instance of the use of habits. Conditioned emotional reactions do not require any prior activity of associated motivation in order that the conditioned stimulus may call out the habitual conditioned emotional response.

(5). Hull's principle *that any drive tends to sensitize all habits* seems to be another gross over-generalization. He cites only one experiment which directly supports the proposition—M. H. Elliott's demonstration that, after some rats had been trained to run a maze for a water reward, they ran almost as well when their motivation was first shifted from thirst to hunger. He also cites, as indirect evidence, four studies that show that habits based on food reinforcement are used to some extent even when hunger has been satiated, and one study by Stone that showed that sexual activity continued in male rats sometimes for eight months after castration.

This is slender evidence to justify such a sweeping conclusion. For, first, we cannot say that there was no sex motivation merely because sex hormones were eliminated from the rats. Second, the fact that food-getting responses occurred after satiation of hunger does not necessarily prove that some other drive sensitized the habit, but may show merely that the redintegrative processes in those situations may have had a small amount of motivational quality to them as a reflection of earlier trials in which hunger had been present. Third, Elliott's findings may come from the fact that food-seeking activity is similar in many ways to water-seeking activity. According to redintegrative concepts, the rats must be understood as having acquired a knowledge of the maze; and that knowledge could be used in the service of any drive.

All this is quite different from saying that any drive sensitizes all habits. In fact, in order that motives may have a high biological utility, it is quite necessary that different motives should tend to sensitize quite different reaction tendencies.

(6). Hull's concept of *"reactive inhibition" or "need for rest" as a partial explanation of experimental extinction and of reminiscence* seems one of the best ideas in the book. He relates the concept to extinction problems in a masterly way. He might have done well, however, to have suggested that this effect is probably the same as that which Kurt Lewin and his group have designated as the phenomenon of satiation; and he might well have integrated his discussion with some of their material that would have thrown additional light on the nature and conditions of this effect (16). Furthermore, Hull should not have attributed this effect merely to the amount of physical work required in the responses. To posit only this one source of reactive inhibition is a sort of infantile objectivism. Even if one does not share the behavioristic viewpoint, he hates to see it employed as though it ought to lead to such conclusions as that.

(7). Implicitly, Hull assumes *that goal-expectations may be disregarded as a factor that might affect behavior.* He does not say this explicitly, but he proceeds as though this might be assumed. For, when he discusses the influence on learning of the nature and amount of the reward, Hull cites experiments in which different groups were compared that were equated as regards motivation (strength of hunger, e.g.), but that were given different amounts of reward. But naturally, as he himself makes clear elsewhere, the curves secured in these studies cannot directly be taken as curves of learning, but are curves of performance. Before it could be assumed that the amounts of learning had been different under different conditions of reward during learning, the goal-expectations ought to have been equated by some further training in which all subjects would have had a chance to learn to expect the same amount of goal-material; and *then* their records should have been compared. Hull has recognized the need for this sort of control when he mentions the question of whether differences of motivation produce differences of rate of learning; but he neglected it on this comparable matter.

G.

A comparison of the above discussion with Hull's Chapter XX, "General Summary and Conclusions," will justify the statement, I believe, that this discussion has examined virtually every one of Hull's main principles of learning and of habit-use. And, amazing though it may seem, what we discover is that there is hardly a principle in the whole collection that can stand up under careful criticism.

This does not necessarily imply that the book has been a failure. As Hull (pp. 398-399) says near the end of the book:

> The primary task of a science is the early and economical discovery of its basic laws. In the view of the scientifically sophisticated, to make an incorrect guess whose error is easily detected should be no disgrace; scientific discovery is in part a trial-and-error process, and such a process cannot occur without erroneous as well as successful trials.

Hull has served this function. He has sharply defined the implications of our common peripheralistic approach to problems of learning, he has worked out some of the implications or predictions of peripheralistic principles, and he has, accordingly, given us a much clearer means of passing judgment on this general mode of approach to problems of learning.

After having thus clearly stated many experimental implications, however, Hull's book has broken down because it has violated so many of Hull's excellent ideas of what constitutes a science. It has been content, particularly, to leave many basic terms undefined, as though they could be selfexplanatory. It has had more faith in broad deductive propositions than it has had in experimental evidence. It has supported its conclusions by consistently ignoring or only vaguely mentioning the evidence that did not harmonize with the conclusions. A truly scientific worker simply does not proceed in this way. He guards most jealously the evidence which fails to fit his formulations. He knows that, out of such discrepancies, even when they appear insignificant, come some of the most remarkable scientific advances—just as, in chemistry, the discrepancies of atomic weight led to the discovery of the isotopes of lead, and this in turn led to the discovery of transmutations of elements.

In some ways it is not surprising that psychologists have not been inclined to have such faith in experimental evidence. We have often had the experience of finding that some initial finding suggested some brave new principle, but then have found with more careful work that chance factors had accounted for the seemingly noteworthy results or that closer analysis showed a series of other factors that might have produced the result originally attributed just to one influence. So, when psychologists are faced with data comparable to the discrepant measurements of atomic weights, they maybe tend too strongly to feel, "Oh, well, there must be something wrong with the data, because that general theory is too well supported." So, we stick to our traditional notions, which tend to be based on the idea that the most scientific understanding of psychological processes will be found in simple, mechanical interpretations. Hull differs from our more common treatment merely in that he shows a somewhat greater faith in such deductive ideas.

But that old approach is getting boresome. Why would it not be possible, instead, to combine Hull's respect for theoretical formulations with much more of wholesome respect for homely factual material? A theoretical system is exciting when it is an integration of the best of scientific observation; when it is a substitute for that, it is wearisome.

Perhaps it is sound practice for us American psychologists to give a central rôle to the general theory of learning. If this practice is to be fruitful, however, our theory of learning must be something alive, something sensitive to relevant experimental evidence. Hence, the important point in the above discussion is not that Hull's theory of learning is open to such serious criticisms, but that, if his theory of learning is weak, the basic theory of learning now accepted by the great majority of American texts in psychology is almost equally weak. Most of such books accepted a fundamentally similar peripheralistic theory of learning, and differ from Hull mainly only in that they have not worked out as explicitly the assumptions and implications of their points of view. The experiments on which they depend are either the same experiments or are comparable to the experiments on which Hull depends. The experiments which Hull has either neglected or refused to consider, with such fatal consequences for the soundness and adequacy of his theory, are the same experiments that are neglected by most other books in psychology. They have no better grounds than has Hull for clinging to principles that center around the concepts that learning is a relatively simple, mechanical, stimulus-determined process and that habits can be described as relatively simple connections between peripheral events. In the early days of objective psychology, such a peripheralistic theory of learning was a promising suggestion. Even now it apparently still carries a nostalgic emotional appeal for many psychologists, perhaps chiefly because of its value as an affirmation of faith in the idea of the availability of psychological processes for scientific study. But as a synthesis of experimental material now available, the peripheralistc approach has outlived its usefulness.

It probably is true that psychology is needed as something that can play a vital social function in the thought of today. But if psychology is to play such a vital social rôle, it must have theories that have some daring, that have enough youth and vitality to take advantage of experimental work even when it challenges traditional lines of thinking. Such re-interpreting is hard—even painful sometimes. But, after all, it's both our job and our privilege.

REFERENCES

1. BERNSTEIN, A. L. Temporal factors in the formation of conditioned eyelid reactions in human subjects. *J. Gen. Psychol.*, 1934, **10**, 173-197.
2. BROGDEN, W. J. Sensory pre-conditioning. *J. Exper. Psychol.*, 1939, **25**, 323-332.
3. GUTHRIE, E. R. The Psychology of Human Conflict. New York: Harper, 1938. (P. 18.)
4. HANEY, G. W. The effect of familiarity on maze performance of albino rats. *Univ. Calif Publ. Psychol.*, 1931, **4**, 319-333.
5. HERB, F. H. Latent learning: Non-reward followed by food in blinds. *J. Comp. Psychol.*, 1940, **29**, 247-256
6. HULL, C. L. Principles of Behavior. New York: Appleton-Century, 1943. Pp. 422.
7. ———. The conflicting psychologies of learning,—a way out. *Psychol. Rev.*, 1935, **42**, 491-516.
8. ———. Mind, mechanism, and adaptive behavior. *Psychol. Rev.*, 1937, **44**, 1-32.
9. ———. Differential habituation to internal stimuli in the albino rat. *J. Comp. Psychol.*, 1933, **16**, 255-273.
10. HULL, C. L., & BASS, M. J. The irradiation of a tactile conditioned reflex in man. *J. Comp. Psychol.*, 1934, **17**, 47-65.
11. HUMPHREYS, L. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. Exper. Psychol.*, 1939, **25**, 141-158.

12. ——— Extinction of conditioned psychogalvanic responses following two conditions of reinforcement *J. Exper Psychol.*, 1940, **27**, 71-75

13 Jenkins, I The postulate of an impoverished reality *Psychol Bull*, 1942, **39**, 580.

14. Kluver, H Behavior Mechanisms in Monkeys Chicago. Univ Chicago Press, 1933. (Pp 258-303)

15. ——— Re-examination of implement-using behavior in a Cebus monkey after an interval of three years. *Acta. Psychol.*, 1937, **2**, 347-397

16 Kounin, J. S Experimental studies of rigidity *Charac. & Personal*, 1941, **9**, 251-272, 273-282

17. Krechevsky, I. An experimental investigation of the principle of proximity in the visual perception of the rat. *J. Exper. Psychol*, 1938, **22**, 497-523

18 ——— A study of the continuity of the problem-solving process *Psychol. Rev*, 1938, **45**, 107-133

19 Lashley, K. S An examination of the "continuity theory" as applied to discriminative learning *J. Gen. Psychol*, 1942, **26**, 241-265

20. Liggett, J R An experimental study of the olfactory sensitivity of the white rat *Genet. Psychol. Monog.*, 1928, **3**, 1-64

21. McGeoch, J Psychology of Human Learning New York Longmans, Green, 1942

22 Razran, G. *Conditioned responses: An experimental study and a theoretical analysis* *Arch of Psychol*, 1935, No 191 Pp 124

23. Ruch, F Psychology and Life New York· Scott, Foresman, 1941 (P 295.)

24 Sackett, R S The effect of strength of drive at the time of extinction upon resistance to extinction in rats *J Comp Psychol*, 1939, **27**, 411-431.

25 Spence, K. The nature of discrimination learning in animals. *Psychol. Rev* 1936, **43**, 427-449.

26 ———. Continuous versus non-continuous interpretations of discrimination learning *Psychol Rev*, 1940, **47**, 271-288

27 Spence, K, & Lippitt, R. "Latent" learning of a simple maze problem with relevant needs satiated *Psychol Bull*, 1940, **37**, 429

28. Swann, H. G. The function of the brain in olfaction I Olfactory discrimination and an apparatus for its test *J Comp Psychol.* 1933, **15**, 229-241

29. Thorndike, E L. Human Nature and the Social Order. New York Macmillan, 1940

30. Tolman, E C Purposive Behavior in Animals and Men. New York· Century, 1932.

31 White, R. The case for the Tolman-Lewin interpretation of learning. *Psychol. Rev.*, 1943, **50**, 157-186

32. Yum, K S An experimental test of the law of assimilation. *J. Exper. Psychol*, 1931, **14**, 68-82

Department of Psychology
University of Oregon
Eugene, Oregon

Footnote from p 7 Cf Sigmund Koch's statement in "Hull's *Principles of Behavior*. A special review" (Psychol Bull. 1944, 41, 269-286): "—Hull's new book—represents the most potent blow thus far struck against—the complementary disease cluster of literary, programmatic, verbal or speculative theorizing independently of—sometimes in spite of—empirical evidence.—*Principles of Behavior* is one of the most important books published in psychology during the twentieth century" (p 269).

The Journal of Genetic Psychology, 1944, 65, 53-66

FACTORS INFLUENCING THE EFFICACY OF PUNISHMENT AND REWARD: THE OPPORTUNITY FOR IMMEDIATE REVIEW, AND SPECIAL INSTRUCTIONS REGARDING THE EXPECTED ROLE OF PUNISHMENT*

Department of Education of The Johns Hopkins University, and Baltimore City College

---

J. M. Stephens and J. A. Baer[1]

Immediate review or inner repetition may readily occur when a subject has answered a question and has found out whether he is right or wrong. This review may consist of some simple reiteration such as, "Minneapolis is *not* the capital of Minnesota," or, "Boston *is* the most populous state capital in the Union," or it may be much more elaborate. It may include self-imprecation or jubilance or a hasty search for cues which may help in the next encounter with the question

It is natural to suppose that one should profit from such review. It should help one to avoid, in the future, any answer which has turned out to be wrong It should help one to retain for the future an answer which has proved to be correct.

The known facts bearing on the influence of such review may be summarized as follows:

1. When there is little or no opportunity for review, the influence of punishment and reward depends on the initial strength of the association. Punishment has considerable influence on strong associations but little influence on weak associations Reward, on the other hand, greatly strengthens weak associations but has little effect on associations that are already strong. Apparently the influence of effects is largely confined to these "Strong but Wrong" and "Weak but Right" conditions (1)

2 When there is an unlimited opportunity for review it is in these "Strong but Wrong" and "Weak but Right" situations that the influence of review shows itself. This fact is illustrated by the comparison of data from two earlier experiments (2, 3) shown in Table 1.

---

*Received in the Editorial Office on October 7, 1942

[1] This investigation is one of a series made possible by a grant to the senior author from the National Research Council In this, as in previous investigations, we are indebted to Dr Edward E Weglein, Superintendent of Baltimore Schools, and Dr Philip H Edwards, Principal of City College, for their cooperation

TABLE 1

THE PERCENTAGE OF ASSOCIATIONS PERSISTING AFTER THE APPLICATION OF SYMBOLIC PUNISHMENT AND REWARD

	Effect applied			Effect applied		
	Right	None	Wrong	Right	None	Wrong
Unrestricted review	86	65	51	59	36	30
Restricted review	82	67	61	53	39	33
Difference	4	2	—10	6	— 3	— 3

In Table 1 the greatest difference between the restricted and unrestricted review is found in the case of the "Strong but Wrong" and of the "Weak but Right" associations

3. Under conditions of unrestricted review students spend a disproportionate amount of time in reviewing "Strong but Wrong" and "Weak but Right" associations (4).

It is apparent that the conditions in which review is most effective are also the (*a*) conditions in which punishment and reward are most effective, and (*b*) the conditions on which students will spend most review when given the opportunity This combination of facts suggests two alternative hypotheses.

1. *First hypothesis* It may be that the "Strong but Wrong" and "Weak but Right" conditions are the only conditions in which effects work anyhow and that the increased opportunity for review merely accentuates these forces which are already in operation. It would seem natural that if any such positive force were at work, an increase in the time during which it is applied would increase its effectiveness If, on the other hand, there were no real force acting, an increase in time would not bring about any increased effectiveness.

2 *Second hypothesis.* It is possible that increased opportunity for review would enhance *any condition to which it is applied* but that, under conditions of free review, the subjects spend almost all their time in reviewing the "Strong but Wrong" and "Weak but Right" associations. The experiment cited (4) shows that a subject who is free to spend as much or as little time as he wishes in reviewing any question, will spend more of his time on these strong hunches *which turned out to be wrong* and the feeble guesses which turned out to be right

The second hypothesis implies that by inducing subjects to spend time in reviewing the "Weak but Wrong" and the "Strong and Right" associations we could enhance the influence of these conditions also. That is to say, that the strong associations which turned out to be right would become even stronger if the subject were to spend some time in reminding himself that this is the right answer to this question Conversely, the weak and erroneous guess would become even weaker if the subject spent some time in considering his error

The present experiment was designed to test this implication of the second hypothesis. The experiment was arranged so as to control the time available for review after *each* question In this way it was intended to rule out the possibility that subjects were spending almost all their review time on some questions and very little on others. After some questions there was minimal or zero chance for review before the next question was presented After other questions a period of four seconds intervened between the application of the effect and the presentation of the next question In still other cases a period of 8 to 12 seconds intervened

There is no opportunity, of course, that all the time provided for review is utilized in review. During this time, however, the question was present on a screen, and the corrected choice was also present There was no possibility of looking at the next question, and little incentive to do anything else but review

MATERIALS AND PROCEDURE

The general procedure was similar to that described in an earlier report (1) From material used in previous experiments and from other sources, questions were selected so as to provide a variety of strength-effect conditions. Some of these questions had been answered with great confidence by previous subjects Other questions had elicited less confident answers, while still others had elicited outright guesses Questions were so chosen that for each degree of strength the most probable answer to some questions would be right, while the most probable answer to others would be wrong The following samples illustrate the various types of questions and show the responsive-effect condition which previous experiments had led us to expect.

Staring at a man's back will
- (*a*) Cause him to turn around
- (*b*) Have no effect on him.
- (*c*) Cause him to feel uneasy

In this question most students chose (*b*), stated that they were sure they were right and found out that they were right.

Our climate during the last 30 or 40 years shows,

(*a*) Less severe winters.
(*b*) More severe winters.
(*c*) No persistent change

Most students chose (*a*) with confidence and found the answer wrong.

The number most frequently associated with "air" is
(*a*) 4
(*b*) 9
(*c*) 12

In this case no choice was predictable Most answers were given as guesses. In some of these weakest questions all choices were considered right, while in others all were considered wrong. In questions with a definite content, however, all rights and wrongs were genuine.

Questions of intermediate strength were chosen from foreign languages and from scientific facts

In all, 60 such questions[2] were chosen and printed on lantern slides Immediately below each question appeared the three choices in a given order. Immediately below that the same choices were given again but in a different order, and below that again the choices appeared in a third order On each presentation two groups of choices were masked out. A different mask was used on each presentation so that the choices never appeared twice in the same order.

The data obtained from 119 High School boys, the material being presented to four different groups made up of from 29 to 37 boys each. The experiment was carried on in a regular class room The boys sat at ordinary school desks facing a projection screen Each student had a pad of answer papers, one paper for each question A sample of the first sheet appears in Table 2

TABLE 2

	Guess	Question 00 Hunch	Sure
a	1 ( )	2 ( )	3 ( )
b	4 ( )	5 ( )	6 ( )
c	7 ( )	8 ( )	9 ( )

In this matrix the words "Guess," "Hunch," and "Sure" refer to degrees of certainty, while the letters (*a*), (*b*), and (*c*) refer to the choices on the screen. By placing a single mark in the appropriate parenthesis the student could indicate both his choice and his degree of certainty. The numbers beside the parentheses were used for coding by the punch card operator. The students were told of the purpose of these numbers and were advised to ignore them.

---

[2] Actually 100 questions were presented to some groups, but the first 20 and last 20 were intended for another aspect of the investigation and were not considered in this paper

The questions were presented once a day on three successive days.

First day's procedure. Question 00 was presented on the screen. The experimenter allowed a certain time for reading and then said "Choice." At that signal each student was supposed to indicate his choice. Immediately after saying "Choice," the experimenter called out, "Question 01. Check with your answer sheet." Upon this signal, the boys turned over the page to Question 01 while an assistant flashed Question 01 on the screen. The procedure was then repeated.

Second day's procedure. The answer sheets to be used on this day had been treated with invisible ink in the manner described elsewhere (1). Students made their marks with moisteners filled with a developer. Contact with the developer caused a symbol to appear in the parentheses. A preliminary question was provided to give some familiarity with the technique. After making a choice for this sample question, the students received the following instructions:

> If an *R* appears from the invisible ink it means that you are right. If a *W* appears it means that you are wrong. Sometimes no symbol will appear. That means that you can't tell whether you are right or wrong. It does not mean that you are right or that you are wrong, or partly right or partly wrong. It means that you just don't know. At other times the sign "&" may appear. This means just the same as when nothing appears, that is, you don't know whether you are right or wrong.
>
> You will be given the same test tomorrow and you should be able to use the information you get today on tomorrow's test.

For two of the four groups additional special instructions were given. These appear below in the discussion of the rôle of the special instructions.

The students were allowed to fill in all nine parentheses of the sample question as practice and to allow some of the novelty to wear off. They were then cautioned that in other questions they must not mark more than one space per question.

On this second day varying periods of delay were inserted between the time of making the choice for one question and the presentation of the next question. The delay periods used were 0 (or as close to 0 as possible), 4 seconds, 8 seconds and 12 seconds. (In treating the data the 8- and 12-

second delay periods are grouped together) For three groups the minimum delay occurred for the Questions 00-19, the four seconds delay for Questions 20-39, and the longest delay for Questions 40-59 For the fourth group this order was reversed. During any period of delay the question just answered remained on the screen.

Third day's procedure On this day no information was provided and no delay periods were introduced. The questions were flashed on the screen and the subjects indicated their choice on the answer sheet, degree of certainty being ignored Following this presentation, other information was secured, but this has no bearing on the topic of the present paper

TREATMENT OF THE DATA

Hollerith punch card equipment was used for the many sorting and tabulating operations

The first problem was to determine the strength of the association just prior to the application of the effects on the second day. For this determination both the consistency of the choice and the student's estimate of his certainty were used The data were sorted into five arbitrary categories as indicated in Table 3

TABLE 3

Degree of Strength	
1	Choices appearing for the first time on Day 2 and labeled "Guess" on that day
2	Choices appearing for the first time on Day 2 but labeled "Hunch" or "Sure" on that day
3	Choices consistent on both days but labeled "Guess" on one day
4	Choices consistent on both days and labeled either "Hunch" on both days or "Hunch" on one day and "Sure" on the other
5	Choices consistent and "Sure" on both days

Although these gradations of strength of choice were available, it was found that when the choices were divided into three delay groups and into five strength groups, there were too few choices in a group. Consequently for this part of the investigation Categories 1 and 2 were combined to form a group of weak associations and Categories 4 and 5 were combined to form a strong group Category 3 was not used.

The influence of increased opportunity for review is shown in Table 4

The trends appear most clearly in the last three figures of each row. The

TABLE 4

THE NUMBER AND PER CENT OF STRONG AND OF WEAK CHOICES PERSISTING TO THE THIRD DAY AFTER THE APPLICATION OF PUNISHMENT AND REWARD WITH DIFFERENT OPPORTUNITIES FOR REVIEW

Strength of choice	Effect Applied	Total number of choices			Number persisting			Per cent persisting		
		Opportunity for review (in seconds)			Opportunity for review (in seconds)			Opportunity for review (in seconds)		
		0	4	8-12	0	4	8-12	0	4	8-12
Strong	Right	314	304	292	281	276	262	90	91	90
Strong	None	309	344	320	229	280	237	74	81	74
Strong	Wrong	264	327	295	119	142	112	45	44	38
Weak	Right	234	228	216	121	136	133	52	60	62
Weak	None	414	439	417	168	184	187	41	42	45
Weak	Wrong	342	361	318	101	121	102	30	34	32

trends for "None" (no information) have no especial bearing on either of the hypotheses.

There is nothing in either hypothesis to lead us to expect a choice followed by no information to be less likely or more likely to persist just because the subject has had time to consider his answer The data for the "No Information" conditions have been included chiefly because of their peculiar trend in the case of the strong associations Here we find no difference between 0 delay and 8-12 seconds delay, but we do find greater persistence in those control choices which were followed by four seconds delay The peculiar relation is not readily ascribed to chance since χ^2 for that part of the table is 6.4 ($P = 05$—) and since the same trend appears in each of two independent sub-groups. At present we do not know how to explain such a result nor how to take it into account in the interpretation of the rest of the study

It is with the remaining four trends that the two hypotheses are concerned. The first hypothesis holds (*a*) that delay will increase the effectiveness of the "Strong but Wrong" situations and the "Weak but Right" situations, but that (*b*) it will not increase the effectivness of the "Strong and Right" or the "Weak and Wrong" combinations The second hypothesis contends, on the other hand, that the effectiveness in all conditions will be increased by an increased opportunity for review.

In determining the dependability of any trend we have used χ^2 (for some trends) and have also calculated results separately for two distinct sub-groups. The results for the sub-groups are shown in Figure 1

The information in Table 4 and in Figure 1 does not support the second hypothesis. In the figures for the total groups and in the graphs of the sub-groups it appears that increased opportunity for review does not increase the influence of "Right" on strong associations. Neither is there any indication that review increases the influence of "Wrong" on weak associations The *apparent* trend in this latter case suggests that review *decreases* the effectiveness of "Wrong" The trend is not entirely consistent from sub-group to sub-group, however, and it is safest to conclude merely that there is no definite increase in the effectiveness of "Wrong" on weak associations.

The data just discussed automatically offer some support to the first hypothesis It is still necessary to show, however, that "Strong but Wrong" and "Weak but Right" associations do profit from review In the case of the "Strong but Wrong" associations there is a slight trend toward less persistence as the review time increases, suggesting the greater effectiveness of

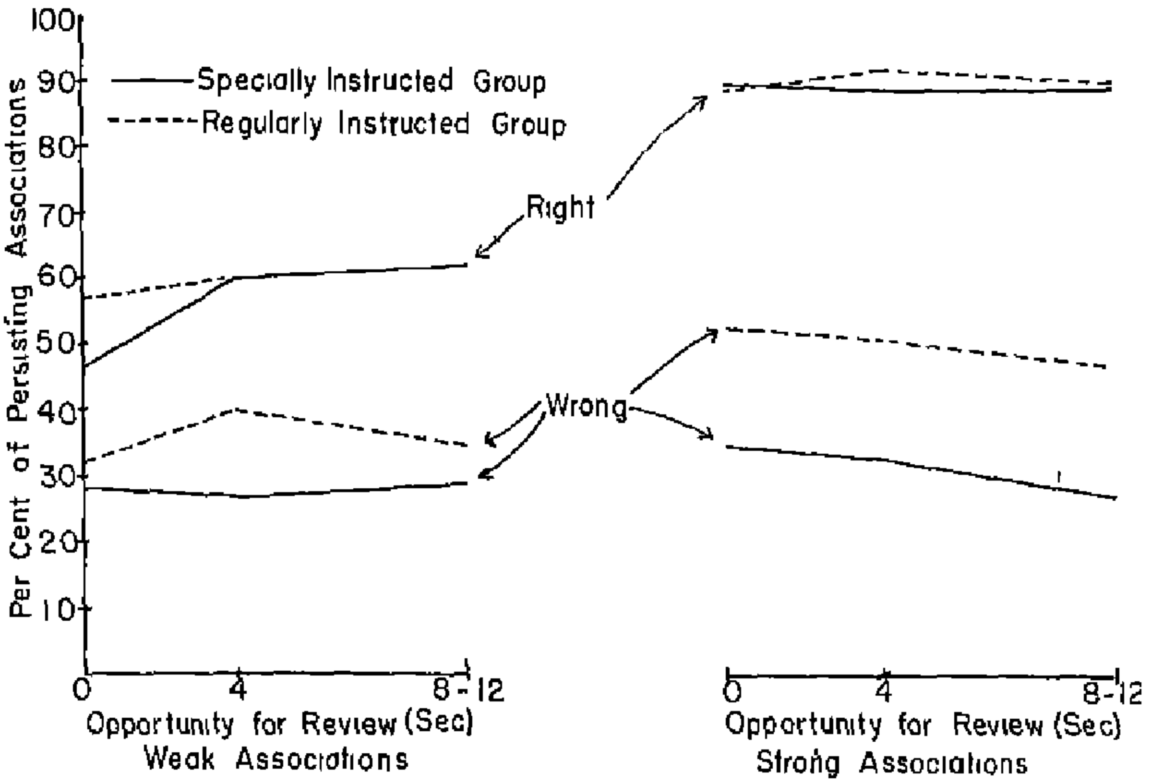


FIGURE 1

THE INFLUENCE OF OPPORTUNITY FOR REVIEW ON THE EFFICACY OF PUNISHMENT AND REWARD WHEN APPLIED TO WEAK AND STRONG ASSOCIATIONS

"Wrong" with greater time for review. This slight trend appears in the graphs for both sub-groups (The absolute difference in these graphs has no importance in this phase of the investigation It is the trend of each group that concerns us) For the total group this trend has a χ^2 of 3 1 (the data being assembled in a 2x2 table and the persistence after 8-12 seconds delay being compared with persistence after 0-4 seconds delay.) This just falls short of significance at the 5 per cent level In view of its approach to significance, and of its appearance in two independent samples, one should judge that this trend has fair dependability.

When "Right" is applied to weak associations its influence seems to increase with increased opportunity for review. This trend is not equally pronounced in the graphs for the two sub-groups but it is in the same direction in each sub-group The χ^2 for the total group (obtained from comparing persistence at 0 delay with persistence for all other delay conditions combined) is 6 8 This is significant at the 1 per cent level

Apparently the previously observed enhancement of effects by review in "Strong but Wrong" and in "Weak but Right" situations, comes from the fact that these are the only situations susceptible to such enhancement and not from the fact that they secure an undue proportion of the review time.

The Rôle of Special Instructions

The special instructions were designed to offset a possible mind-set which may interfere with the optimum functioning of punishment Though a student would ordinarily strive to retain a correct aniswer, he may see little point in trying to get rid of a wrong answer when his new choice is just as likely to be wrong as right. Consequently a student may not consider the elimination of errors as an important end in itself. In our experiment, however, the successful avoiding of an error is all that is expected of punishment If the student avoids the error, punishment is considered effective, no matter whether the new choice is the correct choice or merely another error

Tilton (5) had been concerned with this same problem. In an experiment involving only weak associations he told some subjects that there was one response to each question which must be avoided at all costs while any one of the three responses would be considered correct. Tilton found that subjects so warned were no more successful in avoiding weak wrong choices than the subjects who were, presumably, concentrating only on getting the right answer

Our experiment is not strictly comparable to Tilton's since we used strong as well as weak associations and since our attack on the problem of mind-set was more direct. We merely tried to show the subjects how they could utilize a statement of "Wrong." Toward this end two of our four groups of subjects were given the following special instructions in addition to the regular instructions.

> Remember, it is just as important tomorrow to avoid an answer which is wrong today as it is to choose an answer which is right today If you choose a wrong answer today you still don't know which is the *right* answer, but you do know which is the *wrong* answer and you should be able to avoid that answer tomorrow.

Results

Some evidence of the different performance of the specially instructed groups appear in Figure 1. We see that for the wrong associations the percentage of persisting choices is, in general, lower for the specially instructed than for the regularly instructed.

If we arrange finer gradations of initial strength (and ignore differences

TABLE 5

THE NUMBER AND PER CENT OF CHOICES OF DIFFERENT STRENGTH PERSISTING TO THE THIRD DAY AFTER THE APPLICATION OF PUNISHMENT AND REWARD TO GROUPS GIVEN SPECIAL INSTRUCTIONS AS TO THE EXPECTED INFLUENCE OF PUNISHMENT AND TO GROUPS GIVEN ONLY THE REGULAR INSTRUCTIONS

		Total number of choices			Number persisting			Per cent persisting		
		Strength of choice			Strength of choice			Strength of choice		
Effect applied	Nature of instructions	Weak	Med	Str.	Weak	Med	Str	Weak	Med	Str
Right	Regular	360	454	244	213	350	232	59	77	95
Right	Special	318	474	161	177	376	155	56	79	96
None	Regular	655	707	179	284	440	154	43	62	86
None	Special	615	739	158	255	464	132	41	63	84
Wrong	Regular	486	536	177	174	253	77	36	47	44
Wrong	Special	525	530	125	150	186	33	29	35	26

in opportunity for review) we get the results indicated in Table 5 and in the graphs to the left in Figure 2 Here we have three strength groups made up of a weak group (Categories 1 and 2 combined), a medium group (Categories 3 and 4) and a strong group (Category 5).

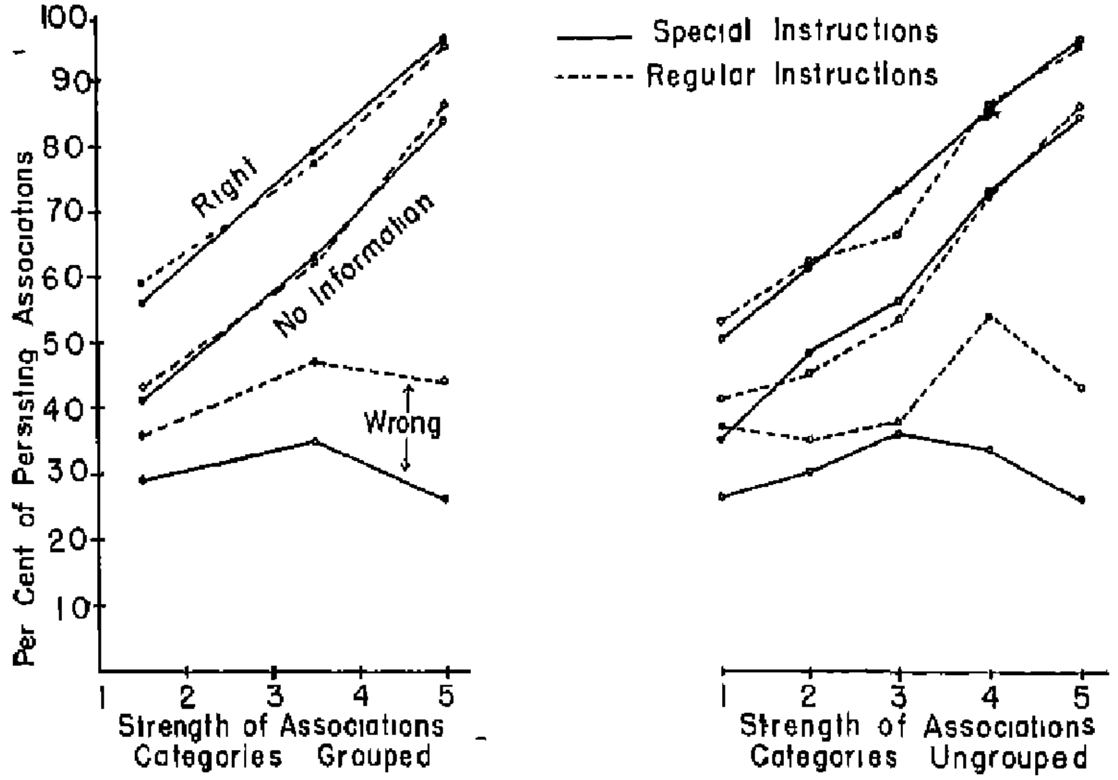


FIGURE 2

THE INFLUENCE OF PUNISHMENT AS AFFECTED BY SPECIAL INSTRUCTIONS REGARDING ITS EXPECTED INFLUENCE

The results show that whereas the groups are very comparable as to the influence of "Right" and of "No Information," they differ markedly in the influence of "Wrong" For each degree of strength the specially instructed group is more successful in avoiding the choices which turned out to be wrong, than is the regularly instructed group

If, in a search for even finer gradations of strength, we plot the results for each category of strength we get the graph which appears to the right in Figure 2 Here again we see that the results for "Right" and for "No Information" are fairly comparable for each group In the case of "Wrong" the specially instructed group is still more successful in avoiding the wrong choices. We notice, however, that the graph for the regularly instructed group is very erratic with the consequence that for some degrees of strength the difference between the groups is very slight

There is no obvious explanation for the peculiar trend of the wrong associations of the regularly instructed group In the previous experiment (1) the graph of the regularly instructed groups was not at all like that of the regularly instructed group in this experiment. It was, in fact, much more like the graph of the specially instructed group of this experiment.

Whatever may be the cause of the vagaries of this peculiar graph, it seems clear that in the general trend, the specially instructed group excelled the regularly instructed group in avoiding associations which turned out to be wrong

SUMMARY

It has already been shown that symbolic punishment weakens strong associations but has little effect on weak associations Conversely, symbolic reward strengthens weak associations but has little influence on strong associations This selective influence of punishment and reward is accentuated when subjects are free to review each choice after the punishment or reward has been applied and are also free to go back to any previous choice and review its correctness.

These facts suggest that the amount of time spent in review may play a causal rôle in the selective action of punishment and reward Due to the element of surprise, a subject may spend more time in reviewing a strong hunch which turned out to be wrong than in reviewing a strong hunch which turned out to be right

In the present experiment the subjects were given no choice as to the amount of time spent in review. When opportunity for review was extended equally for all conditions it was found to enhance only the influence of punishment on strong associations and the influence of reward on weak associations. It is suggested that these conditions profit by review because they are the only conditions in which effects function and not because they secure an undue proportion of the review time

The investigation shows that subjects become more likely to profit from symbolic punishment when they are told that it is just as important to avoid an incorrect response as to retain a correct response.

REFERENCES

1 STEPHENS, J M. The influence of symbolic punishment and reward upon strong and upon weak associations *J Gen Psychol*, 1941, **25**, 177-185

2 ——— Further notes on punishment and reward. *J Genet Psychol*, 1934, **44**, 464-472.

3. STEPHENS, J M, & BAER, J. A The influence of punishment on learning when the opportunity for inner repetition is reduced. *J Genet Psychol.*, 1937, **51**, 209-217
4. STEPHENS, J. M, BAER, J. A & STINNETT, B J Spontaneous delay following the application of punishment and reward to strong and weak connections. *J Genet Psychol*, 1940, **57**, 279-288.
5. TILTON, J W The effect of "Right" and "Wrong" upon the learning of nonsense syllables in multiple choice arrangement. *J. Educ Psychol*, 1939, **30**, 95-115

Department of Education
The Johns Hopkins University
Baltimore, Maryland

Baltimore City College
Baltimore, Maryland

The Journal of Genetic Psychology, 1944, **65**, 67-88

CONGENITAL ALEXIA AND ITS RELATION TO OPTIC PERCEPTION*

Bellevue Hospital, New York City

PAUL SCHILDER, PH.D., M D[1]

The practical importance of congenital alexia or congenital word blindness has been appreciated especially by Hinshelwood in England, Warburg in Germany, and Orton in this country. Monroe has made an extensive tudy of cases of reading difficulties in children and has compared the normal and retarded children. Orton has developed an interesting theory concerning the genesis of congenital word blindness, especially emphasizing the nterference of left and right cortical impulses.

But the theoretical importance of the so-called congenital word blindness in its relation to the present aphasia discussion has not yet been considered fully. Investigation on seven cases have been made on the nature of ongenital word blindness from the point of view of its position in the system of aphasias and agnosias. The examination consisted of presentation of letters in orrect and reversed position, of numbers, and of short and long words and sentences The letters and the words used were both printed and written. The examination also included letters, numbers, words, short sentences, writen spontaneously and from dictation, and the copying of printed and handwritten words and sentences. Pictures were presented to the subject and a description elicited. Letters, words, numbers, and figures used by Orton in his est were presented to every subject. Pictures, drawings, words, numbers, figures, and all the other material used by Orton were also exposed tachistoscopically. The subjects were not only asked to report what they saw but also, in many instances, were asked to reproduce the material by drawing or writing. The pictures varied from relatively simple outline figures of houses, trees, heads, to somewhat complicated landscapes, one of them having a cluster of trees, a church spire, and a road The numbers and letters used were the three-quarter inch, gummed letters used in making posters Some written words were also included. Subjects were also examined in the perception of movements exposed to them for a very short time The tachistoscope was a Shutter model, making an exposure of about 1/35 of a second, or longer when so desired Several examinations were given to each subject, as the children tended to get tired after 30 or 40 minutes of work.

---

*Dr. Paul Schilder died on December 8, 1940 This paper was written in 1935, but was transmitted to this JOURNAL in 1942 by Dr. Lauretta Bender Received in the Editorial Office on October 15, 1942

[1] Acknowledgment is made to Dr. Elaine F. Kinder, for her assistance in the examination of the patients studied, and to the Psychology Laboratory at Columbia University for the loan of the tachistoscope which was used for presentation of the material

This paper covers the study of seven cases of specific reading disability. They were brought to light in connection with the routine psychological examination on the psychiatric service at Bellevue Hospital. They were either ward patients in the hospital, or Mental Hygiene Clinic patients who could be admitted to one of the wards for a brief period of observation. Six were boys and one a girl The youngest was eight years and two months old, the oldest 14 years of age Two of the boys were colored, one born in the United States, the other in Porto Rico.

Case 1 — J S.

This 12-year-old white boy was sent in by the Children's Society because of a poor school record. His Mental Age on the Stanford-Binet scale was 10 years, 5 months; on the Army Performance scale 11 years, 9 months, and on the Healy Picture Completion 2 was 9 years, 5 months. With special reading material he showed no difficulty in the recognition of letters and in *copying* There was *considerable difficulty in recognition of longer words*, but he had learned to read simple words and sentences fairly well. He had no difficulty with numbers and his copying was invariably correct. Tachistoscopic exposures were recognized readily unless the stimulus was a rather long or unfamiliar word. With simple and familiar words no trouble was encountered. He tried to read longer words by sounding the letters Often he would sound one letter and then make errors in the adjoining letters. There were almost no reversed letters Words printed "mirror-wise" were recognized as "turned-around" and were read with great difficulty. It was easier for him to assemble words than letters There were occasional optic mistakes (Table 1)

When "1" is read instead of "I," the difficulty is apparently in optic perceptions, also when "doy" is read instead of "dog," the trouble lies in the perception of the single letter. It is the same type of difficulty which mistakes "M" for "N." Otherwise, it was remarkable that this boy succeeded in the synthesis of only the first part of long words and his separation of the word "Beat" into two halves

TABLE 1

Test Material		
Presented in print	*Read*	
At	Correct	
You	Correct	
Long	Correct	
In Brooklyn	Correct	
News of	News of — News of	
Local Interest	Inter —	
Territory	Terr—Tart—	
Dictated	*Written*	
Sixty-nine, one, two, three	Correct	
(own name)	Correct	
I have a little dog	I have a litter dog	
I like to go to school	I like to go to sch-so-st-school	
We are going to have some candy	We are going to have some car-cac-kn-can	
Dictated	*Spelled verbally*	
Dog	D-O— I had it a minute ago I had it on my mind I can't say it It isn't W, it isn't C	
Little	L-I-T-R-E . I left out the E	
Dog	I got it now D-O-G	
Tachistoscopic Exposure (Print)	*Read*	*Written*
¶	P backwards	Correct
D	Correct	Correct
Walter	W-E-, That was some word I couldn't get it A-L-K-E, a couple of E's	
(Second exposure)	W-A-E-R	
(Without tachistoscope)	Correct "That wasn't the same word"	
Beat	B-E-E-T, B-E-T-T B-E-T	Bett, Bat

Case 2: — P. F

This was a 13-year-old boy with a mild Froelich syndrome (pituitary infantilism) with pudgy hands and features, soft skin, undescended testicles, and a BMR of plus 6 He scored a Mental Age of 9 years, 6 months and an *IQ* of 70 on the Stanford-Binet Scale, and a Mental Age of 10 years, 2 months (*IQ* 75) on the Army Performance Scale. The problem was inability to learn throughout childhood with poor power of concentration, poor memory, and difficulties in reading and writing

With the special reading material, he showed marked difficulty in the reading of simple words. He read reversed letters as readily as those correctly printed and recognized mirror words at least as quickly or possibly faster than ordinary words There was, however, hesitation and slowness in the identification of mirror-written words as such He had no difficulty in spelling words, but there was a marked hesitation and inability in putting letters together into words The patient was unable to read words that he did not know. He wrote from left to right, and in work with anagrams used the correct order of placing letters. There was marked difficulty in writing words from dictation, but copying was consistently correct.

This patient made many more optic mistakes in single letters than did others of the group studied. He made similar mistakes as the others in the reading of words Errors with the tachistoscope were as frequent as when words were presented directly.

Unlike the other patients studied, this boy had considerable difficulty with numbers, which he freqeuntly misread when shown to him tachistoscopically The errors often seemed quite unrelated, as for example, 181 was read as 840 and 373 as 180. He had, however, no trouble in writing numbers from dictation (Table 2).

This patient, too, made mistakes between *B* and *D, N, A, M, H* and *D*. He also copied *A* instead of *S* It may be supposed that these are optic difficulties, especially as there are also mistakes in copying During other test periods he made mistakes between *b* and *d*, *q* and *g*, *p* and *d*, and showed several instances of mirror reading. It is remarkable how he was able to substitute for an unknown word one that he knew, even when there was not more than one letter in common with the word offered.

TABLE 2

Test Material		
Presented in print	*Read*	*Spelled*
SEE	ESS-SEE	
ASS	Correct	
NO	ONE	N-O
ON	Correct	
SAW	"I don't know"— ASS	S-A-W
WAS	HIS	W-A-S
DO	Correct	
BOY	DOG	D-O-G
UP	"Don't know what the letter is"	U-P
800	Correct	
BAT	"Don't know—BUT"	B-A-T

TABLE 2 (continued)

Test Material		
(Second test period)		
DOG	Correct	
YOU	"Can't read it"	Y-O-Q-U
AT	Correct	
SEE	Correct	
ME	"Can't say it"	M-E
ON	Correct	
BEAT	"Can't say it—DEAT"	D-E-A-T
HEM	DEM	Correct
MAMA	MAN	M-A-N-A
ASS	SEE	M-S-S
404	400	404
373	373	
1551	155-1	
649	6-40-6	
800	80	
Dictated	*Written*	
CAT	Correct	
DOG	Correct	
LITTLE	eltli	
BIRD	Bef	
ON	No	
At	Atn	
SEE	Correct	
ASS	af	
GOOD	go	
DOG	Correct	
ON	am	
WEATHER	of	
MAN	anam	
MAMA, BEAT, ON, and COST were presented in print and copied correctly		
Tachistoscopic exposure (Print)	*Read*	
ON	ON—I am not quite sure	
Cost	C—COE	
181	181	
DOG	DOY, BOY	
ME	Me, Spells M-E	
SEE	Correct	
AT	E, A-T, at	
You	Correct	
Walter	W-EA	

CASE 3: — W. B.

An eight-year-old negro boy was referred to the clinic for admission to the hospital as a behavior problem He had a congenital infantile paralysis of the right arm and leg His Mental Age, as measured by the Stanford-Binet Intelligence Test, was 6 years, 6 months. His vision was 16/25 in each eye.

When presented with the special reading material (Table 3), he displayed no difficulty in the recognition of letters, in spontaneous writing, copying, tachistoscopic reading of letters, numbers, or short words How-

TABLE 3

Test Material

Presented in Print	*Read*	*Written* (R = right hand, L = left hand, arrows indicate direction of writing)
I GO	L-G-O, that's No	
ᗡO	D-O, that's No	
WALTER	W-A-L-T-E-R (after some hesitation)—Ah ha! That's my name I knew you were going to fool me It's Walter	
NO		NO (R) ← O WO (L) ←
SEE	Eyes, (spelled S-E-E-E)	SEE (L) → ƎƎS (R) ←
DOG	D-O-G "Is that dog? Yes sir, it is"	DOG (L) → ƆOᗡ (R) ←
ᗡOG	"Oh, it's dog again, only this time it's this way"	ᗡOƆ (L) → ƆOᗡ (R) ←
GO	That's OG	ƆO (L) ← ƆO (R) ←
WALTER	That's my name (Starts printing from right to left, then writes from left to right)	W-A-L-T-E-R (L)
GO	GO—now it's O G	GO OG
OD	D-O	D-O →
3	3	3
61	Nineteen	19 ←
41	Fourteen	18 ←
12	Twelve	12 →

TABLE 3 (continued)

	Test Material	
Tachistoscopic exposure (Print)		
YOU	You I only seen U, I didn't see the rest	U U O → U O ←
BOY	V-O-Y— B-U-O-Y cannot read it	B O Y →
ME	E—cannot read it	M E →
MAMA	Does not know	M
DOG	DOG, D-O-G	DOG →
ASS	S E E E, Eyes (when repeated, he says "Again S E E E, of course")	
373		E—, then 37—, "I wonder what's the other one?" — 373.
38	Thirty-eight	38 ←
81	Does not know	8 1 1 →
181	Does not know	1 8 1 ←
181	One-eight-one	1 8 1 →
699	Six-nine-nine	9 9 6

ever, he was unable to read long words or to write them from dictation He recognized reversed letters. He made frequent mirror mistakes when writing spontaneously and showed a marked tendency to write from right to left This tendency was as marked in the writing of numbers and work with anagrams as it was in his writing of letters and words. These mirror mistakes appeared only in spontaneous work or during dictation, but were not present during the copying of words or letters

Single words were almost always spelled correctly but could not be read. The tachistoscopic performance was somewhat worse than when the words were presented during the usual time span. The recognition of objects, pictures, and movement was uniformly good, whether the stimuli were presented tachistoscopically or not. The right and left confusion was confined to reading and writing Practically all writing was done with the left hand, due to the right-handed hemiplegia Right-hand writing was helped by and determined by the left hand. Left-hand writing was done with the arm encircling the paper above the writing.

Anagram Dictation. He was given cut-out letters and asked to pick out the correct letters and to form the word DOG While he spelled vocally "D-O-G," he joined, at the same, the letters "G-O- ," and then reversed, the entire word. For CAT he selected C and T, spelled C-A-T, found *A* and placed it correctly For RAT he placed R-T, spelled R-A-T, and was satisfied with incomplete anagram. UP, BOY, and NO were joined correctly He seemed to have little difficulties in the orientation of letters after the first word. He always began with the first letter, spelled the word orally, and worked out the problem by finding the letters consecutively. But when told to place the letters B-A-T or H-A-T, he could not read the found word.

Case 4. — A. S

An 11-year-old white boy was referred to the clinic by an Orphan Home, where he had been for 8 years, with the report that he was annoying and quarrelsome in school and unable to keep up with his grade work. His Mental Age was 8 years, 10 months on the Binet Scale, 8 years (Median

TABLE 4

Test Material	
Presented in print	*Read*
Once	O-N-C-E (cannot read)
there	T-H-E-R-E (cannot read)
was	W-A-S—Where
little	L-I-T-T-L-E (cannot read)
so	so
CANDY	C-A-N-D-Y (cannot read it) (Do you want to have it?) "Yes, sir." (Is it a dog, cat, elephant?) "No." (Is it candy?) "Yes."
THIEF	(cannot read it) (Is it thirst, thunder, Thursday, murderer, house, or thief?) "It is thief."
As	As
Cat	Cat
Candy	(Reads now correctly)
TABLE	Cannot read it (Is it clothes, shoe, ear, eye, threatening?) "No" (Is it a chair?) "Yes." (Are you sure?) "Yes."
1284	Twelve-hundred-and eighty-four

TABLE 5 (continued)

Test Material		
Tachistoscopic exposure (Print)	*Read*	*Write*
CANDY	Correct	
CAT	Correct	
Monkey	"I don't know how to spell it I don't know what it is."	
Third exposure	Monkey	M-O-N-K
Ass	Correct	
THIEF	Thie	
Shoes	Stockings	S-E-T
LION	Correct (What is the last letter?) "N"	L-O
Stick	Stockings	
Frog	"Don't know what it is or how to write it."	
Anagram dictation	*Puts together*	*Spelled verbally*
ALBERT	Correct	
ON	O-N-G	
DOG	GOD	D-O-G
HOUSE	HAST	H-A-S-T, reads "House"
ME	Correct	
Cow	KIAT	K-I-A-T, reads "Cow"
Cat	Correct	
DOG	GOOD, reads "Dog" (What is the first letter?) "G" (Read it) DOG	
Up	Correct	
BOY	BOUT Reads BOY	
GOD	Correct	

Mental Age) on the Pintner-Paterson Performance Test, 10 years, 9 months on the Healy P. C. II, and 12 years on the Porteus Maze. His eyesight was normal

In the work with the special reading material (Table 4) he had difficulties with combination of letters. He recognized individual letters, could pick them out, copy them correctly, and spell short and long words without hesitation. There was slight difficulty with two and three letter words, but with words of more than three letters the patient was completely lost, unless the word happened to be the one he had especially learned. He knew his name, both first and last, could write them from dictation, and recognized them immediately when they were presented in either print or written form Most words of four or five letters, however, were impossible for him and he would respond by saying. *"I don't know,"* or "I can't" When urged, he would attempt to sound out the word, usually getting the first letter correctly without being able to synthesize the entire word For example, "mama," after 8 seconds of effort, was called "man" "Walter," after being spelled and told that this is a boy's name, was read "Wallace." This boy had no difficulties with numbers.

When shown drawings, words, numbers, and designs tachistoscopically there was no difficulty in perception of anything but words, and even the recognition of simple words seemed to be somewhat facilitated by a brief exposure It seemed as though tachistoscopic presentation tended to assist the putting together of the letters into a single word,—if the word presented was not a long one, and was familiar.

Copying either with anagrams or with pencil was done without error, though often the patient could not read what he had just written. There were a few mirror mistakes in the writing of single letters shown tachistoscopically and occasional mirror reversals in the use of anagrams. These, however, were not frequent and formed but a small part of the reading errors. There was no tendency to mirror writing or left handedness.

It is remarkable that the patient occasionally has a vague general idea concerning the contents of words he can't read This is especially outstanding concerning the word "Candy." It is also a remarkable reaction when he chooses the word "Stockings" for "Shoes." One could believe that it is only the similarity in the first letters, but then he selects the word "Chair" for "Table" When he chooses the word "Lung" for "Lion," there is only the identity in the first letter responsible for the choice.

When he has put together a long word in anagram dictation incorrectly, he spells correctly what he has put together, but reads it as the word asked from him.

Case 5

A white boy, age 8 years, 2 months, was referred by the Children's Court with the complaint that he acted like a mental defective He was restless and disobedient on the Ward. His Mental Age was 8 years 8 months on the Stanford-Binet Intelligence Test, scored a Median Mental Age of 10 years on the Pintner-Paterson Performance Scale, and 7 years, 4 months on the Healey P. C. II His eyesight was normal

In the work with the special reading material (Table 5) he was found to have considerable difficulty in the recognition and writing of letters He knew the names of many letters but seemed to encounter trouble in their associa-

TABLE 5

Test Material		
Presented in print	*Read*	*Written*
DOG	No response	Doc, D-O-C
ASS	No response	Acc, A-C-C "It is an S, it is the first letter of my name"
M, R, S, and E were read correctly		
G	C	
P	B, P	
D	D	
ON	DO	ON, O-N
ME	"I don't know the whole word"	N-E
DOG	"It is the same word, I don't understand it good"	D-O-C
37, 811, and 181 were read correctly		
1551	One hundred five hundred fifty one	
18 152	One eight one five two	
Tachistoscopic exposure (Print)	*Read*	*Writes*
ME	"I don't know what it is"	N
2nd Exposure		M-T, N-T
3rd Exposure	Same word I saw it twice"	F
ASS	"It's A"	a-s-s (uses written letters)
NO		To-, N-O, To
UP	Up	P-U
BOY	"There is an 'R' at the beginning"	
2nd Exposure	"I only saw the last letter"	Y
337	7-3-7	7-3-7
2nd Exposure	3-3-3	3-3-3
3rd Exposure	7-3-3	7-3-3
4th Exposure	3-7-3	3-7-3
83	83	
Anagrams	*Put together*	*Read*
BOY	Y-O-ꓭ	
Girl	"L, where is L, where is G?" Puts together F-C, and does not find the rest.	
DOG	C-O-R	D-O-G, Dog
	Anagram copying	
HEN	T-H-E-N, T-H-E-H-E-N	
THE	T-H-	
BAG	Correct	

tion with the appropriate letter forms For example, when working with anagrams and trying to write his name, he would say. "*P*, *P*, where is *P?*", fingering over many of the letters in front of him, passing often again and again over the letter for which he was looking Though this occurred especially frequently with such letters as *D*, *B*, and *P*, it took place also with *W*, *S*, *V*, and other letters of the alphabet

There was much less difficulty with numbers, in fact, he made a few errors with them except in the reading of longer numbers presented tachistoscopically Though there were mirror mistakes in the recognition in reading of letters, and a few mirror mistakes in such words as "on" and "saw," this formed a relatively small percentage of his total errors. Tachistoscopically, he had no difficulty with perception of objects, faces, directions. Even the most complex of the pictures presented were accurately reported He also was able to describe correctly and in detail pictures that were presented non-tachistoscopically. He seemed to have no difficulty in the recognition of directions.

Copying, either with anagrams or with pencil, of short or long words was done without error, but spontaneous writing of words, the reading of either long or short words presented on a card or presented tachistoscopically was impossible for him Numbers were hardly any problem for him. The reading of numbers, writing them spontaneously, or copying, showed no errors, and only a few mistakes were noted for numbers read tachistoscopically.

Of the entire group studied, this boy was the most difficult one to work with. He easily became irritable and often refused to continue. He used his right hand exclusively Tendency toward mirror writing was not noticed. He often read *S* as *C*, *G* as *C*, *P* as *B*, mistakes which point to difficulties in optic perception

CASE 6. — R. T.

A Negro boy of 9 years, 11 months was referred to the clinic because he was disobedient, restless, beyond control He had started school at the age of 6 but could not make progress and had been placed in an ungraded class.

His Mental Age on the Stanford-Binet Scale was 7 years, 6 months, 10 years on the Pintner-Paterson Performance Test, and 8 years, 10 months on the Healey P. C II

During the examination (Table 6) he was alert and responsive when presented with non-verbal material, but became restless and distractable when confronted with verbal material for any length of time In work with the performance material, there was a tendency to use the left hand

During the experiment with the special working material he had no difficulty with spontaneous writing of letters, copying with pencil or anagram letters or short words. There was an occasional difficulty with the recognition of letters, considerable difficulty with short words, and complete inability to handle long words. The difficulty with letters was not noted when letters were presented tachistoscopically. The difficulty with short words was

TABLE 6

Test Material

Presented in print	*Read*	*Spelled*	*Copied*
as	at	correct	correct
it	it	correct	correct
left	little	correct	correct
ball	black	D-A-L-L	dall
pod	dog	P-O-D	God
balk	black	B-A-L-K	dalk
pig	pirl	P-I-G	dig
plea	(refused)		glea
bed	(refused)		ded
plea	(refused)		clea
Boy, Cat, Man were read correctly			
The boy has	The boy has		
a dog	a ball		
HAS	his	has	
HORSE	Hees		
THE	There, the		
WAGON	woke		
Tachistoscopic exposure (print)	*Read*		
YOU	Y-O-T		
GO	correct		
BAT	B-O-L		
HEM	H- I did not see it so good		
ASS	A- What's the other?		
2nd Exposure	A-B-S, bass		
69	correct		
800	correct		
482	23		
Anagram dictation	*Put together as*		*Spelled*
GO	og		
BOY	bog		
G	correct		
Y	correct		
19	92		
BOY	dog		
354	3054		
TODAY	todog		D-O-G reads today I can't spell it
weather	water		
24			
61			
192	0042		
354	3054		
Has	us		
on	mot		

present but seemed slightly as marked. His trouble with long words was unaffected by tachistoscopic presentation.

This boy had greater difficulty with numbers, both in copying and in writing from dictation, than most of the cases studied.

He wrote mirror-wise with his left hand and normally with his right hand with nearly equal facility. His errors were of the same sort and as frequent in his left-hand as in his right-hand writing. In spontaneous dictation and in copying of print, he changed *b* to *d*, *p* to *d* and *p* into *g* There was trouble with his spelling, but he was able to overcome this easier. In reading he read correctly only words that he knew, and changed unknown words into known words. He was unable to put single words together. The letters were spelled correctly but he could not put them together unless he knew the word The difficulty with right and left appeared only in the single letters There was no tendency to mix the order of the letters.

CASE 7 — C M

A 14-year-old white girl was referred to the clinic because she was unable to concentrate or remember anything pertaining to school work, but was handy in manual activities. Her Mental Age on the Stanford-Binet Scale was 9 years, 4 months, 15 years on the Pintner-Paterson Performance Test, and 10 years, 10 months on the Healey P. C. II.

In the work with the special reading material she was able to read single letters and numbers without difficulty Simple words were read slowly in spelling fashion There were many errors in writing. For example, Brooklyn was written Brooklny, Little was written Litter In writing from dictation words were omitted or written with great difficulty and with frequent mistakes When letters and numbers were shown tachistoscopically, she seemed to have no difficulty in recognition on the whole, though there were occasional mirror mistakes such as reading a capital *E* as *3*, or reading the word *no* as *on*. Longer words were not read at all, though they were recognized as words Relatively long numbers were read correctly Recognition of pictures was excellent. This patient showed the same general tendencies that the others did, as far as the examination went Examination was not completed because she was taken to Chicago by her mother.

SUMMARY

In general, all of the cases studied showed the characteristics reported by Orton in his cases of word-blindness. Letters of the alphabet, either printed or written, were generally well recognized. There were numerous mistakes between *b* and *d* and *B* and *D*, *g* and *p*, *Y* and *G*; *p* and *d*, *M* and *N*; *m* and *n*; *b* and *g*, *A* and *S*; *d* and *t*; *C* and *S*, *A* and *M*; *I* and *L*, *T* and *L*; *H* and *D*. It is true that these mistakes were much more outspoken when the word was presented as a whole, and rather rarely occurred when the single letter was presented The same mistakes were made in spontaneous writing or in writing from dictation There were fewer mistakes concerning Arabic numerals. Spelling of presented words was generally good and only those mistakes occurred which appeared also in reading of single letters. The difficulties varied in the different cases There was also a variation of mirror mistakes, and was conspicuous only in two cases. The tendency to sound one or two of the initial letters of a word and then jump at the rest was most marked in two cases where there had been the greatest amount of school training

The following tendencies were noted.

1 In practically every case the difficulty was confined to letters and words Four digit numbers were recognized and reported correctly where four letter words were not recognized at all Such four digits could be even read as a continuous number, as for instance, "twelve-hundred and eighty four " Of the seven cases studied only one showed difficulties in the recognition of numbers, but there was less hesitation than in the recognition of words

2 Mistakes concerning the Orton pictures were rare and only outspoken in Case 3, which showed the greatest amount of mirror mistakes. Otherwise, the recognition of pictures was very good, even when the details of the pictures was rather complicated Tachistoscopic presentation of pictures revealed prompt recognition of the parts of the pictures. Within the limits of the tachistoscopic perception of normals, the pictures were perceived as a whole in practically every case. We never observed mirror mistakes in the permanent or tachistoscopic presentation of pictures The accuracy of picture recognition was in marked contrast with the difficulty in word recognition

3. On the whole tachistoscopically presented words were as readily recognized as permanent exposures. Familiar words seemed to be, at times, more readily recognized tachistoscopically. One can say that the tachistoscopic presentation did not increase the difficulty.

4 Difficulties with reading words, spelling or writing words was present in all cases studied, even when the single letters and sounds did not offer any trouble. Noted tendencies toward mirror writing and reading were found in two cases. In the others, mirror writing was entirely absent. Mirror mistakes in reading were either absent or played a very unimportant part quantitatively

5. The optic perception of *movements and their directions* was excellent, even when the movements were exposed only in the tachistoscope.

6. All of our patients could differentiate well between their left and right side, could perform movements with either hand according to order, and were also able to touch the right part of their bodies with the left hand and vice versa, according to given orders.

DISCUSSION

Orton emphasizes the importance of mirror mistakes for the genesis of reading and spelling disabilities He supposes that the process of learning to read entails the elision from the focus of attention of the confusing memory-images of the non-dominant hemisphere which are in reversed form and order, and the selection of those which are correctly oriented and in correct sequence. According to this theory, in reading disability there *has been an incomplete elision of the memory patterns in the non-dominant* hemisphere, and therefore left or right sequence may be followed in attempting to compare presented stimuli with memory images. This leads to confusion or delay in selection "Our analysis points to confusion because of reversals in the memory images of symbols, resulting in a failure of association between the visual presented stimulus and its concept." He points out that the children with these confusions see as others do, but fail to learn to elide completely one of the two anti-tropic anagrams registered as a pattern for a later comparison, which forms the basis of recognition. Orton distinguishes between static reversals which occur in the single letter, for instance, *in the mistake between b and d,* and between *p* and *q,* and kinetic reversals, where there is a tendency to assemble the letters in sinistrad or alternating sinistrad or dextrad order. In reading disability, this tendency may exhibit itself in only a few letters of a word, in a syllable, or in a whole word. Orton is inclined to explain in this way all changes in the order of letters. In the in and out reversals, the child may spell in both directions and *how* becomes *wow,* and *rag* spelled as *raggag.* A good instance of this occurred in Case 2 where *little* was written as *eltth.*

Cases 3 and 6 of our study showed indeed a marked left and right disorientation, but in the majority of our cases the left and right mistakes did not play an important part and may be completely absent, as in Case 1. This makes it at least improbable that Orton's theory covers the whole field. It is at least worthwhile to discuss whether even extreme difficulties in the left and right orientation of letters and words can produce the enormous difficulties the children have to coordinate in reading the sounds of two letters which are as such very well known to them One would at least expect that when a child can read *u* and *p*, that he should be able to make one word out of these sounds! In our Case 3, the boy copies correctly "*m-e*," but still is not able to make a word out of it It is hard to believe that difficulties of this kind should be merely secondary to the difficulties in the right and left orientation of the letters. However, Orton is certainly right when he supposes that the difficulty in visualization of words by mirror mistakes may drive the child to phonetic spelling and phonetic mistakes in spelling and writing

We find, in addition, in cases which do not show mirror mistakes, difficulties in reading single letters which can hardly be explained in another way than in the primary difficulty in the optic perception or imagination of letters. Here belong the mistakes between *L* and *I*, and *T* and *L* or *g* and *y*, or in another case between *C* and *S* I believe, therefore, that there exists primary optic difficulties in the perception and use of letters, at least in some cases Probably the mistake between *M* and *N*, and *m* and *n* belongs in this category, although definite proof was not possible But there is no question that even the optic mistakes in these cases are very different from those one finds in the cases of so-called pure alexia in adult persons, in which, at least in one group of cases, not only the wrong letter is perceived, but also the form of the letter is distorted. The condition of the eyes is not responsible for such mistakes We observed trouble in the refraction only in Case 3 In Case 2, there was some impairment of the eyesight in connection with a hypophyseal condition But even in those cases the impairment of the eyesight is certainly not responsible for the difficulty in the perception of letters Case 2, by the way, is the only case in which there is some difficulty in the perception of numbers, more than simply reversals

Certainly, the optic mistakes concerning the letters cannot be taken as the cause of the reading disability. They are merely an indication that we deal with the primary disturbances in the perceptual optic sphere, which are not merely due to a difficulty in the dominance between left and right.

But in all these cases one difficulty was outstanding, and that was the difficulty of coördinating the sounds which are more or less correctly connected with the single letter, with the sequence of letters in the re-write word This inability was especially evident in short words such as *me, up, was, bat,* and *hem* It is hard to see how this inability, which occurred in reading as well as in writing, could be secondary to any other disturbance, either in the optic sphere alone, or in the left and right correlation. When severe cases of that kind are able to *read the whole word it is due to the* fact that they have learned the word as a whole, but they are absolutely unable to analyze the word Even the attempt to analysis causes great disturbance This seems to be the reason while these words are sometimes pretty well perceived in tachistoscopic exposure. Here is also the explanation for the fact that sometimes only an identity of one more or less important letter causes misreading of one word for another. One child read *stockings* instead of *shoe*. Sometimes, as in Case 4, one gets the impression that there is a vague knowledge concerning the concept of words which he cannot read. All this leads to the conclusion that we deal with a primary disturbance concerning the sound structure in the written words. This disturbance can be compared with the difficulty with the spoken word as observed in some aphasias In some aphasias we find that sounds are not put in the right place, that they change their positions in the whole of the word Reduplication takes place. We are not justified to believe that all these mistakes of sequences in sounds are due to interferences between the right and left hemispheres, although interferences of that kind may also be of importance in aphasia The lesion of a center of the left hemisphere certainly pushes the function back to a level of lower integration. The disintegration between left and right impulses may be one part in this disturbance.

We know that the destruction of the left practic region may liberate *mirror tendencies in the left hand and even improve the left-hand mirror* writing. But the destruction of the left region, which alters the dominance, *changes at the same time the inner development of the action, generally, and* the writing especially. The change in the dominance of the hemisphere, although *important in itself, is not sufficient to explain all* the other disturbances The German literature on Aphasia has repeatedly tried to explain the paraphasias of the so-called sensory aphasia as the expression of the activity of the right brain In the so-called sympathic apraxia, which occurs on the left side when either the left practic region or the connection between the left and right practic region (corpus callosum) is impaired, we find the motor activities of a more primitive type. But even in this instance the removal of the dominance of the left hemisphere provokes a type of reaction which is identical with the lower type of integration which results from every lesion of a cortical center of high order and is not merely due to the tendency of mirror reversal

In the congenital reading disability there are sufficient hints in the so-called optic mistakes and in the inability to co-ordinate the single sounds with the whole word and to differentiate the whole word into the single sounds and in their written expression that we deal with a primary inability and insufficiency of centers This primary inability may release the mirror tendencies. This mirror tendency certainly can have a new disturbing effect on the process of reading, writing, and spelling, but the mirror tendency in itself is insufficient to explain the whole picture We come, therefore, to the formulation that the congenital reading disability is due to an incomplete function of centers It is the inability of the patient to differentiate the spoken word into its sounds, and to put sounds together into a word It makes that word to have meaning only as a whole. The sound has a meaning of its own. Single sounds are brought in connection with a written letter, but the written letter cannot be put together into a whole The written word, even when connected with the object named, cannot be separated into its single sounds and letters The letters correlate with the sound within the word This formulation agrees pretty closely with the formulations of Monroe, Illing, and Bachmann. It differs from the formulation of Orton only in so far as Orton has the opinion that this disturbance, which he emphasizes too, is secondary to the disturbances in the dominance of the hemispheres, which increase the mirror tendencies, whereas according to our opinion the disturbance is primary and the release of the mirror tendencies secondary. The released mirror tendencies (in so far as we agree with Orton) may increase the difficulty. It is worthwhile to give some consideration to Case 3 from this point of view This child had, according to the history, not only a birth paralysis of the right arm but also a subsequent encephalitis with right side hemiplegia. One is astonished to see that in this case the mirror tendencies have remained outstanding and that static and kinetic reversals are in the foreground. But even in this case the difficulties in putting together the sounds into a word are so great that the mirror tendencies do not explain the whole difficulty. We do not deny that in such a case mirror tendencies are a factor of enormous importance One could be tempted to range this case with the rare cases in which congenital reading difficulties are connected with organic lesions (cases by Gunther, Currier, and Dewar). In our case as well as in the cases of the literature, it is doubtful whether the focal lesion provokes the reading difficulty or is not merely a factor which adds in reading disabilities which are based upon the insufficient development of those centers which guarantee the development of the faculty of reading Orton has rightly emphasized the familial character of reading disabilities. Potzl, in this connection, has pointed to the investigation of Karplus who found that the configuration of the gyri in the brain may show hereditary and familial characteristics We do not doubt that the congenital reading disability can and must be interpreted from a neurological point of view. We, therefore, consider the opinion of Gate's that reading disability is due to faulty habits in learning as contradictory to the facts

We deal with a trouble of the agnostic type Bachmann's attempt to separate it from agnosia and aphasia is not justified. We agree with Orton and Monroe who point out the fact that the intelligence of these cases of congenital reading difficulties is good, or at least sufficient for the acquisition of reading and writing We certainly do not deal with a general lack of intelligence. All of our cases, with the exception of Case 2, were sent to us as behavior problems, which resulted to a large degree from the inefficiency in reading We emphasize again that we deal with a rather isolated gnostic disturbance.

It is worth-while to remember how far different congenital reading disability is from pure word-blindness which results in adults from a lesion of the gyrus-angularis (according to Pötzl gyrus-lingualis). Cases of that type are very often unable to recognize letters. In the less serious cases *the difficulty is restricted to the recognition of words But they are unable* to copy letters and words, whereas the spontaneous writing is good They are only unable to read what they have written themselves. The pure word-blindness is a mere optic product. In the congenital reading disability, the difficulty concerns the inner structure of the word and its sounds. It is an agnostic trouble, not so much concerning the merely optic sphere, but a sphere which is nearer to the intellectual life than the optic perception. It is true that every gnostic function is also an intellectual one, but the intellectual function which is disturbed in the congenital reading disability cases is of a higher level than the function which is disturbed in the cases of pure word-blindness. We deal here with an isolated trouble of gnostic intellectual function

This formulation strictly contradicts tendencies which are almost predominant in modern neurology. Head and Goldstein do not believe that isolated disturbances of gnostic and speech function can occur by cortical lesion. A disturbance which impairs one gnostic function damages, according to their opinions, every gnostic function Lashley has come to very similar formulations in experiments on rats Orton's and our patients did not show any aphasic trouble. There was no general intellectual impairment. Our study placed special emphasis on the careful examination of optic perception, which disclosed no pathology or functional inability. Deviations from the norm were not observed during tachistoscopic expositions of Objects We say that careful investigation did not reveal any difficulties in the optic perception If optic perception as such would be disturbed, one would expect that the tachistoscopic exposure of letters and words would impair the perception enormously As Bachmann already has shown, this is not the case The increase in the difficulties by tachistoscopic exposure in the children with reading disabilities was in no case greater than the difficulties a normal child would experience under similar conditions In Case 5, the tachistoscopic perceptions of words seemed to be even better than the perception under ordinary conditions, but this was only true of words familiar to him The tachistoscopic exposures made it easier to take the word as a whole

Our final interpretation would be that the final mastery of the written and spoken word, concerning differentiation and integration of sounds, is a very complicated function These difficulties vary in languages, and increase in English where the sound of a single letter is not identical with the sound in the word The faculty varies in different children (Monroe). These variations are probably due to a different development of those parts of the brain which are indispensable for the process of reading In the serious cases we probably deal with a dysfunction of a cortical apparatus. This dysfunction expresses itself in the integrating and differentiating difficulty in optic mistakes concerning letters and in increased mirror tendencies. The inherent mirror tendencies as well as mirror tendencies originating from other sources, and intellectual difficulties may alter and increase the primary trouble which is an isolated trouble of gnostic intellectual function.

CONCLUSIONS

The basic difficulty in congenital reading disability is the difficulty to differentiate the spoken word into its sounds and to put together the sounds of a word. Words and single sounds are brought into connection with a written word and a written letter, but the written word and the written letter cannot be integrated and differentiated. Two other independent difficulties may be present—mirror mistakes and trouble in the optic perception of letters. There are no troubles in other optic perceptions. We deal with an isolated trouble in a gnostic-intellectual function.

References

1. Bachmann, F Uber Kongenitale Worthblindheit *Abhandl. aus de Neurol, Psychiat, Psychol, und ihren Grenzgeb*, 1927, **40**, 1-72.
2. Currier, F P, Jr, & Dewar, M Difficulties in reading in school children *J Mich State Med Soc*, 1927, **26**, 300-304
3. Gates, A. I Improvement of Reading New York Macmillan, 1929
4. Günther, F Beitrage fur Psychiatre, Pathologie und Klinik der sogenannten Lesechwache. *Z f Kinderfor.*, 1928, **34**
5. Goldstein, K. Die Lokalization in der Grosshirn. *Deut. Z. f. Nervenheilk.*, 1923, **77**, 891
6. Illing, E Über Kongenitale Wortblindheit. *Monat f Psychiat u. Neurol.*, 1929, **71**, 297-355
7. Lashley, K. S Brain Mechanism and Intelligence Chicago Univ Chicago Press, 1929
8. Monroe, M. Children Who Cannot Read Chicago. Univ Chicago Press, 1932
9. Orton, S. T *"Word blindness" in school children* *Arch Neurol & Psychiat.*, 1928, **14**, 581-651
10. ———: Familial occurrence of disorder in the acquisition of language *Eugenics*, 1930, **3**, 1
11. Potzl, O Die Aphasielehre Optischagnostische Storungen In *Handbuch fur Psychiatrie* Vienna Deuticke, 1928

Dr Lauretta Bender
Psychiatric Division
Bellevue Hospital
New York City

The Journal of Genetic Psychology, 1944, 65, 89-96

TEST SCORES FOR CLERICAL APTITUDE AND INTERESTS BEFORE AND AFTER A YEAR OF SCHOOLING*

University of Pennsylvania

Samuel F. Klugman[1]

A. The Problems

In June, 1939, all the students enrolled in the commercial curriculum of the Mastbaum Vocational High School in Philadelphia were given an extensive battery of tests among which were the *Minnesota Vocational Test for Clerical Workers* by Andrew and Paterson and the *Vocational Interest Blank for Women* by Strong. There were 208 white, native-born, female subjects ranging in age from 14 years 5 months to 18 years 9 months, with a mean of 16 years 4 months. In September of the same year, all the newly enrolled pupils were given the same battery. There were 124 white, native-born girls in this group, ranging in age from 13 years 11 months to 16 years 10 months with a mean of 15 years 6 months. In June, 1940, after both groups had received two semesters of schooling, all those who had taken the battery the year before (whether in June or September) were asked to take the clerical aptitude and vocational interest tests again. On this occasion, there were 207 students to whom the tests were administered

As a result of these testing periods, there became available to the present writer data which could be utilized to answer two problems· (*a*) Will the test scores for clerical aptitude and interests be the same after schooling as before, and (*b*) will the relationship between interests and aptitude remain the same after a year of schooling as before?

B The Tests

In this study, the two parts of the Minnesota Vocational Test for Clerical Workers—a name checking and a number checking test—are treated as separate tests in accordance with the manual of directions (2). Of the 17

---

*Received in the Editorial Office on November 4, 1942

[1] The writer wishes to acknowledge his gratitude to Mr Robinson, Principal of the Mastbaum Vocational School and his assistant, Mr Paterson, for granting him permission to work with the students of the school, to Dr Bover, Director of the Division of Educational Research, and his special assistant, Dr Gordon, for arranging for him to test the students. Dr Gordon, too, very kindly read this article and offered several helpful suggestions

occupations for which the Vocational Interest Blank for Women can be scored, the two most appropriate ones—General Office Worker and Stenographer-Secretary—were selected. These, too, are treated as separate tests.

The adequacy of these tests as measures of clerical interest and aptitude can easily be ascertained by recourse to such authorities as Williamson (25), Viteles (24), Bingham (3), Fryer (9), Buros (4), and Paterson, Schneidler, and Williamson (17)

C. RESULTS

1. *Numerical Scores Before and After A Year of Schooling*

Table 1 shows the means, standard deviations, and critical ratios for all tests. It will be noted that in every instance there is a substantial improvement. This gain is so great (20.69 and 16.22 points for the Strong Interest Blank, 14 47 and 12 75 points for the Minnesota Clerical Aptitude test) that in every case the difference is significantly reliable On the basis of these data, a counselor cannot assume that the picture painted by our aptitude and interest test scores will be the same after a year of schooling as before. As a result, Andrew's contention that the Minnesota Test "is relatively free from the influence of clerical training and experience" (1, p. 44) is open to some doubt Similarly, both Bingham's and Strong's claims concerning the stability of interest patterns have not been substantiated by the findings when the scores are treated numerically.

TABLE 1
MEAN, *SD*, AND CRITICAL RATIO BEFORE AND AFTER SCHOOLING

	Mean	*SD*	*Mean*	*SD*	*CR*
General office worker	55.90	37.23	76.59	38 64	5.81
Stenographer-secretary	75 45	35 99	91.67	35 60	4 61
Name checking	100 56	24 03	115.03	24 07	6 27
Number checking	110 17	21 44	122 92	21.65	6 02

2 *Variations in Test Results After A Year of Schooling*

Table 2 gives a detailed picture of the variations in test scores after one academic year of clerical training. In every instance, many more students improved their scores by earning more points than did students whose scores declined. Thus, on the interest test there were two and three times as many students who scored more than did those who scored less. On the aptitude test, the ratio was about five students scoring more to every student scoring less.

TABLE 2
VARIATIONS IN TEST RESULTS AFTER A YEAR OF SCHOOLING

	On retest, scored				Range of deviations		Mean of deviations	
	More	Less	Same	*N*	More	Less	More	Less
General office worker	151	53	3	207	1 to 140	1 to 91	36.40	27.18
Stenog - sec'y	136	68	3	207	1 to 140	1 to 70	35.49	22.07
Name checking	166	32	9	207	1 to 110	1 to 70	20.58	7.31
Number checking	166	36	5	207	1 to 56	1 to 30	19.50	9.08

The deviations from the first testing period to the second not only occurred more frequently in the direction of improved scores, but the range was greater, too. Thus, the range of improvement on the interest test ran from one to 140 points while the range of declined scores ran only as high as 70 and 91 points. Similarly, on the aptitude test, the improved scores deviated as much as 110 points while the declined scores deviated only 30 points at most.

Moreover, not only did more students improve their scores and over a wider range, but the mean of these deviated scores was higher for them than for the students who earned less. On the interest tests, the mean gain was 36.40 and 35.49 for the General Office Worker and Stenographer-Secretary tests, respectively, while the mean loss was 27.18 and 22.07 points. On the aptitude tests, the mean gain was 19.50 for the Number Checking tests, while the mean loss was 9.08 points, the mean gain for Name Checking was 20.58 while the mean loss was 7.31 points.

Further corroboration of these findings is found in Table 3, where it will be noted that fewer letter ratings declined (on the General Office Worker test there were six and on the Stenographer-Secretary test there were ten instances) than did letter rating which improved (23 cases for each of the interest tests).

TABLE 3

STRONG LETTER RATINGS AFTER A YEAR OF SCHOOLING

Rating in 1939	Rating in 1940	General office worker	Stenographer-secretary
A	A	172	170
A	B+	4	6
A	B	1	2
A	B—	0	0
B+	A	16	16
B+	B+	4	3
B+	B	1	1
B+	B—	0	1
B	A	5	7
B	B+	1	0
B	B	2	1
B—	B	1	0
		N=207	207
		$\frac{D}{\sigma D_p}$=10.26	9.88

3. *Strong Interest Letter Ratings After A Year of Schooling*

The Strong numerical scores were converted to letter grades and a comparison was made for the years 1939 and 1940 This was done to check Bingham's statement: "Interest patterns do change with time, but rarely to such an extent that an *A* rating becomes a *C*" (3, p. 74) The results appear in Table 3. The evidence confirms Bingham's claim for out of 207 students who took the same test in 1939 and in 1940, 178 or 86 per cent obtained the same letter rating on both occasions for the General Office Worker test. Likewise, for the Stenographer-Secretary test, 174 or 84 per cent obtained the same letter rating on both occasions. When a critical ratio was established between these percentages and the remainder (all of which showed some degree of change), it was found to be 10.26 for the General Office Worker and 9.88 for the Stenographer-Secretary test results in favor of unchanged letter ratings

These findings, of course, do not contradict or nullify those based on numerical scores They are merely the result of coarser groupings and the unique manner in which the letter grades were distributed at the time of first testing. On the General Office Worker test, for example, 177 students got an *A* rating on the first testing. Therefore, even though a student earned a considerably higher score on a retest, her improvement did not appear because the letter grade *A*, which occurred so overwhelmingly at the time of first testing, was the highest possible rating.

4. *Oldest Versus Youngest Members*

The evidence of a foregoing section indicates that after a year of schooling, there was a reliable increase in scores for each of the tests Is the improvement due to schooling or merely to maturation? In order to answer this question, the scores of the 30 oldest and the 30 youngest members of the group of 207 students at the time of first testing were taken. The youngest students' ages ranged from 13 years 11 months to 15 years 1 month with a mean of 14 years 10 months. The ages for the oldest members ranged from 16 years 8 months to 18 years 9 months and the mean was 17 years 3 months. It will be noted that between the two means there is a difference of 2 years 5 months This is greater than the 9 to 12 months which elapsed between the first and second testings of this study. Furthermore, there is no overlapping between the upper age limit of the youngest students and the lower age limit of the oldest students If maturation is responsible for the improvement which has occurred, then the older group will show reliably higher means than will the younger group If there should be found no significant difference, it may be assumed that increase in age is not the reason for the improvement but rather the schooling which the students received during the academic year

Means, sigmas, and the critical ratios between the two groups for all four tests were determined and the results appear in Table 4 It will be seen that in no instance is there a reliable difference As a result of these findings, it is probable that there is a schooling effect which tends to improve scores on the clerical interest and aptitude tests used in this study

TABLE 4

MEAN *SD* AND CRITICAL RATIO BETWEEN YOUNGEST AND OLDEST MEMBERS OF THE GROUP ($N=30$ each)

	Youngest		Oldest		
	Mean	*SD*	*Mean*	*SD*	*CR*
General office worker	58 86	34 86	55 06	37 86	40
Stenographer-secretary	80.26	36 99	75 14	35 78	55
Number checking	110 67	23 62	117 33	16.26	1.27
Name checking	105 67	22 87	110 00	18.73	80

5. *Coefficients of Correlation*

The scores earned by the 207 students on all four tests in 1939 were correlated with the scores earned on the same tests in 1940. The results appear in Table 5. Following Garrett (11, p 342), the coefficients for the interest test scores indicate substantial relationship and those for the aptitude test scores denote high relation.

TABLE 5

COEFFICIENTS OF CORRELATION BASED ON SCORES BEFORE AND AFTER SCHOOLING

General office worker	579 ± 031
Stenographer-secretary	483 ± 036
Number checking	.733 ± 021
Name checking	.765 ± .018

Not only were correlation coefficients between the numerical interest scores of 1939 and 1940 obtained by the product-moment technique, but coefficients of contingency were also obtained by employing the converted letter grades Since the C's were based on a 4 × 4 fold classification they were corrected for broad categories (as per Garrett, pp. 390-391). For the General Office Worker scores a corrected C of 597 was obtained while for the Stenographer-Secretary scores a corrected C of .559 was found. It will be noted that these contingency coefficients agree closely with the product-moment ones Following Lindquist (16, pp 214-216), the reliability of the differences was determined for both interest tests and they were found to be not significant—the t value between C and r was .233 for the General Office Worker and for the Stenographer-Secretary it was 1.043. "The observed difference would have to be approximately 2.576 times its standard error to be significant at the 1 per cent level, or about 1 960 times its standard error to be significant at the 5 per cent level" (16, p. 216).

D. SUMMARY AND CONCLUSIONS

In an effort to determine whether an individual's clerical aptitude and interest test scores remain stable in spite of an academic year of schooling, 332 white, female, native-born students enrolled in the commercial curriculum of a vocational high school were given, among others, a clerical aptitude and a vocational interest test. Of this group, 208 had had from one to five semesters of schooling; the other 124 were new enrollees One year later (by which time 125 students had dropped out of the curriculum), after they had had two terms of schooling, the students were tested again with the same aptitude and interest tests (same forms) with the following results.

1. On all four tests—(*a*) Name Checking, (*b*) Number Checking, (*c*)

Stenographer-Secretary, and (*d*) General Office Worker—the mean gain for each was statistically significant.

2. When the results were broken down, it was found for all tests that more persons improved their scores than did those whose scores declined or remained stationary, improved scores showed a wider range than did the declined scores; and the means of the improved scores were higher than those of the declined ones.

3. Bingham's statement that letter ratings of *A* on the Strong interest test rarely becomes a *C* was upheld However, this is not to be interpreted as a refutation of the results based on numerical scores It merely indicates that, due to the high percentage of students who obtained the highest possible rating of *A*, no increase in score could be discerned even though it was present

4 When the scores of the 30 oldest and the 30 youngest members of the group were compared on all tests, in no instance was there a reliable difference between the means. This may be taken to mean that it is not maturation which is responsible for the improvement in scores from 1939 to 1940, but probably schooling.

5. When the scores earned on each test in 1939 were correlated with those earned on the same tests in 1940, it was found that a high relationship existed for the clerical aptitude test scores and a substantial relationship for the clerical interest test scores. On the interest tests, coefficients of contingency based on letter grades agreed closely with product-moment coefficients based on numerical scores.

REFERENCES

1 ANDREW, D M Analysis of the Minnesota Vocational Test for Clerical Workers. *J. Appl Psychol.,* 1937, **21**, 18-47

2. ANDREW, D M, & PATERSON, D G Manual of Directions, the Minnesota Vocational Test for Clerical Workers New York Psychological Corporation

3 BINGHAM, W. V Aptitudes and Aptitude Testing. New York Harper, 1937.

4. BUROS, O The Nineteen Forty Mental Measurements Yearbook. New Brunswick, N. J

5. COWDERY, K M. Measurement of Professional Attitudes Differences between Lawyers, Physicians, and Engineers *J Pers Res*, 1926, **5**, 131-141

6 CRAIG, D R. The Preference-Interest Questionnaire in selecting retail saleswomen *J. Pers Res* 1924, **3**, 366-374

7 FREYD, M Selection of Promotion Salesmen. *J. Pers Res*, 1926, **5**, 142-156

8. FRYER, D Predicting Abilities from Interests. *J Appl. Psychol.*, 1927, **11**, 212-225

9. ———. The Measurement of Interests New York Holt, 1931

10. GARRETSON, O. K Relationships between Expressed Preferences and Curricular Abilities of Ninth Grade Boys New York Columbia University, 1930 Pp. 72.

11. GARRETT, H E Statistics in Psychology and Education. New York: Longmans, Green, 1939.

12. HARTMAN, R, & DASHIELL, J. F An experiment to determine the relation of interests to abilities *Psychol Bull*, 1919, **16**, 259-262

13. HUBBARD, R. M. A measurement of mechanical interests. *J. Genet Psychol*, 1928, **35**, 229-252.

14 KITSON, H D Interest as a criterion in vocational guidance *Educ Rev*, 1916, **52**, 349-356

15 KORNHAUSER, A W. Results from a quantitative questionnaire of likes and dislikes used with a group of college freshmen *J. Appl. Psychol*, 1927, **11**, 85-94

16. LINDQUIST, E F *Statistical Analysis in Educational Research* New York. Houghton Mifflin, 1940.

17. PATERSON, D, SCHNEIDLER, G C, & WILLIAMSON, E. G. Student Guidance Techniques. New York: McGraw-Hill, 1938.

18 REAM, M. J Ability to Sell: Its Relation to Certain Aspects of Personality and Experience Baltimore: 1924

19 REMMERS, H. H The measurement of interest differences between students of engineering and of agriculture *J Appl Psychol.*, 1929, **13**, 105-119.

20. STRONG, E K. Interests of engineers: A basis for vocational guidance *Pers J*, 1929, **7**, 441-454.

21 ———. Change of Interests with Age. Stanford University: Stanford Univ Press, 1931

22. ——— Manual for Vocational Interest Blank for Women Stanford University Stanford Univ. Press, 1938

23 THORNDIKE, E. L Early interests. Their permanence and relation to abilities *Sch. & Soc*, 1917, **5**, 178-179

24 VITELES, M S Industrial Psychology New York: Norton, 1932

25 WILLIAMSON, E G. How to Counsel Students. New York McGraw-Hill, 1939.

26 WOODWORTH, R S Dynamic Psychology. New York Columbia Univ Press, 1918.

Medical and Psychological Examining Unit No. 6
Psychological Section
Keesler Field, Mississippi

The Journal of Genetic Psychology, 1944, 65, 97-111

THE MOTIVATIONAL VALUE OF BARRIERS FOR YOUNG CHILDREN*

Department of Psychology, Radcliffe College

IRVIN L. CHILD AND ELIZABETH ADELSHEIM[1]

A. INTRODUCTION

In a monograph entitled *The Influence of Barriers Upon Strength of Motivation* (1), H. F. Wright has suggested as a general principle of behavior the proposition that a barrier enhances a positive valence. Phrased differently, this proposition states that an individual is more likely to strive for any goal if there is a barrier between him and the goal than if the latter is readily accessible While there can be no doubt that this statement holds true of some individuals in some situations, stating it as a general principle makes it at once appear contradictory both to common sense and to psychological knowledge For example, it has frequently been observed that approach to a goal is faster and more energetic as the spatial barrier between the organism and the goal is diminished, these observations have given rise to the generalized term "goal-gradient" It has also been observed in many experimental situations that organisms generally choose the shortest or least effortful of several alternative paths to a goal, a tendency which has led to the formulation of a "law of parsimony" or "law of least effort."

The observational basis for Wright's general principle consists of a group of experiments with children and with adults In general, each experiment involved placing subjects in a situation where they could choose between two desirable objects, one of which was less accessible than the other either because of distance or because of physical or social barriers. The extent to which subjects chose the less accessible object, the specific nature of the objects being experimentally controlled, was taken as an indication of the extent to which the principle was justified that the presence of a barrier increases the valence of a goal.

---

*Received in the Editorial Office on December 4, 1942

[1] The experimental work reported here was carried out entirely by Miss Adelsheim under the direction of Mr Child. Responsibility for the report in its present form is shared by the two authors We are grateful to Dr J G Beebe-Center for having called our attention to Wright's work and the problems it raised.

Wright's experiments with adults seem to demonstrate that under certain conditions a majority of adults will choose the less accessible of two desired objects. The experimental set-up was such that this conclusion is not perfectly certain. On the other hand, there can be no doubt but that people could be so trained that they would make such a choice, and many may have been so trained in the course of an ordinary upbringing in our culture Even if in a majority of situations adults in our society preferred goals which were difficult to obtain, we would not be justified in expanding that observation into a general principle of behavior, for the motivational factors involved are notoriously subject to the influence of training

The experiments with children, if their results were equally in favor of Wright's hypothesis, might come closer to justifying its erection into a general principle It would still remain altogether possible that the preference for objects difficult of attainment was a product of social learning in our society, rather than a general principle of behavior, but the latter possibility would certainly merit serious consideration The results of Wright's experiments with children, however, are considerably more equivocal than those with adults

The type of behavior with which Wright has dealt is of great importance. The preference for, or at least the willingness to engage in, the pursuit of goals which are difficult of achievement, is a major variable in the character of individuals and, from the point of view of the group, a very significant element in morale. It is important to know whether, and how, this preference may be developed in the course of rearing children As a preliminary to further study of this problem, it should be helpful to know whether any credence should be given to Wright's hypothesis that such a preference is a basic or primary phenomenon of behavior. For that reason, it was decided to duplicate some of his experiments with children, to see whether his somewhat hesitant conclusions would stand up under the test of greater numbers of cases.

Wright's work included a number of experiments with young children, using various goal objects, types of barrier, and methods of expressing choice Because of limits of time, only a portion of the experiments could be reproduced In the first place, it was decided not to deal with Wright's "past-barrier" experiments, relating to the effect of a previously existing barrier upon present motivation, for they appear to introduce additional variables difficult to control, such as the tendency to resume an interrupted task We considered the "present-barrier" experiments, relating to the effect of an immediately present barrier, to be more relevant to Wright's main hypothesis. Out of the "present-barrier" experiments, we selected certain ones on the basis of the simplicity of apparatus required, minimal interference with the program of the nursery schools whose cooperation we were enjoying, and apparent relevance to the theoretical question at hand.

B General Procedure

All of our experiments were performed in nursery schools, and all of the experimental work was done by a single person[2] The subjects included both boys and girls, ranging in age from two to eight years, with the majority being between three and five years old[3] Each of the separate experiments was done in entirety on a single day and at a single school. In general, an attempt was made to duplicate the conditions of Wright's original experiments as closely as was possible on the basis of the descriptions given in his monograph For that reason a detailed description of each procedure will not be given and the reader may be referred to Wright's publication. Our procedures differed from Wright's, however, in certain important ways, as follows

1 In explanation of the failure of the subjects under certain conditions to prefer the barred goal object, Wright introduces the concept of generalized object-tension If the child is overwhelmingly devoted merely to obtaining the general type of goal-object that he knows awaits him, so that the presence of two alternative paths and objects is not important to him, Wright suggests, the principle of parsimony becomes effective and the child chooses the immediately available object This generalized object-tension, Wright felt, was especially likely to develop when the child knew ahead of time that the experimenter was going to give him, let us say, a piece of candy In order to eliminate this possibility we have, wherever possible, kept the subjects ignorant of what was awaiting them until the actual moment of the experiment In certain schools it was possible to have each child, after he had acted as a subject, go into a room or yard where he would have no contact with the children who had not yet participated in the experiment In the description of each experiment, we mention whether this control was possible. In all cases, neither the experimenter nor those

---

[2] We are extremely grateful to the directors and staff of the following nursery schools, all in the vicinity of Boston, for their generous cooperation: Frances Stern Nursery School and Kindergarten, Hecht House Nursery School, South End Day Nursery, Sunnyside Day Nursery, and Wheelock Child Center

[3] Our results showed no difference between boys and girls, nor between age groups

in charge of the nursery school gave any information about candies or toys to the children prior to the actual experimental instructions.

2. In order, again, to avoid a generalized object-tension which might encourage the choice of the more accessible object, Wright felt that the two objects between which a choice has to be made should be obviously different, yet they also need to be sufficiently similar so that intrinsic differences between them would not be the primary determiner of the choice. For that reason he used in early experiments two pieces of candy of the same type and approximately equal in size but different in color. He felt that for many subjects the color of the candy appeared to be a very important factor in determining the choice Consequently, he decided it was more desirable to use candy differing only in shape, and this recommendation has been followed in all of the present experiments which involve candy. Throughout, the child was asked to choose between two bonbons which were of about the same size and of exactly the same color but which differed in shape. For the various subjects in a single experiment the shapes were, of course, alternated in a random order as between the available object and the barred object. For the same reason, the toys used as goal-objects in one experiment were all of one color In this respect our experiments follow the procedure of Wright's later experiments but differ, in accordance with his own recommendations, from his earlier experiments.

3. It was not possible to use, as Wright did, rooms specially equipped for the purpose of the experiment, that is, empty of everything except that which was to be used for this purpose. All of our experiments were done in rooms within the nursery school with which the children were completely familiar, rooms into which they had been perhaps daily, and the arrangement of furniture and other objects in the rooms was modified only insofar as was necessary for the experimental set-up Again, it might be expected on the basis of Wright's monograph, that this change would make more likely the choice of the barred object by the children, for one explanation that he gives for the inconclusiveness of many of his experimental findings is that the experimental room was strange to the children, so that the experimental situation was less life-like and more artificial than it should have been.

4 In many of his experiments, Wright used children who had already served as subjects in previous experiments in the series. This circumstance creates considerable difficulty in interpreting the results For example, with a similar experimental set-up, Wright found that when two goal-objects

differed in color a number of children chose the barred one, whereas in a previous experiment in which the objects were identical nearly all the children took the available one He interprets this difference as being due to the degree of identity of the goal-objects. When it is noted that many of the same children took part in the two experiments, however, one has to ask whether the important variable may not instead be a sort of satiation; having previously taken the available object, many of the children may now wish to try the barred one Because of this sort of difficulty, we used each child only once and for a single experiment.[4]

C. THE INDIVIDUAL EXPERIMENTS: PROCEDURES AND RESULTS

1 *Action Choice Experiments*

Wright's experiments were divided into two general groups according to the manner in which the subject was to express his choice. In one set, called "action-choice experiments," the child had to obtain by his own action the goal-object that he preferred In the other group, "verbal-choice experiments," the child was instructed to tell the experimenter which object he preferred and the experimenter obtained it and handed it to him The former will be considered in this section. Throughout, in order to facilitate comparison, we will refer to each separate experiment with the name given by Wright to his corresponding experiment

a String-suspension candy experiment In this experiment two strings were hung in one corner of the room, one in each of the two converging walls 15 inches from the point of convergence, with a piece of candy resting in a triple-pronged hook at the end of each The length of one string was adjusted so that the candy would be 11 inches within the reach of each child, and the length of the other was adjusted so that the candy would be 11 inches outside of each child's reach These two are referred to respectively as the available and the barred object. The lengths of the strings were adjusted to each child's reach, the latter having been measured by the experimenter on a previous day's visit to the school

As in Wright's experiment, a play-room chair was placed three feet out from the corner of the room and the child was made to stand on the chair before he was given instructions, so as to encourage his perceiving the chair as a possible means of overcoming the spatial barrier between him and the less accessible piece of candy. It appears to us from Wright's monograph that he may have had a second chair placed immediately under the barred object, and hence in close proximity to the chair on which the subject stood while receiving instructions We felt that such an arrangement might almost reverse the degree of availability of the two pieces of candy, hence we used only one chair Once the child was standing on the chair, the experimenter stood behind him, asked him to look into the corner of the room (if he was not already doing so), and said, *"Do you see one piece of candy on each side? All right then, you may get either one you want"*

---

[4] We used a total of 140 children, 20 in each of the seven experiments

The room used for the experiment is one in which are kept extra materials which the children place there or procure as the occasion demands One half of the room was cleared for the experiment In this experiment the children who had participated were kept separated from those who had not

Results Of 20 children used as subjects in this experiment, 19 chose the available piece of candy and only one chose the barred piece. In the similar experiment reported by Wright, 10 children chose the barred candy and 9 the available one.

Discussion An overwhelming majority of our subjects chose the available candy, whereas Wright's subjects were about evenly divided between the two alternatives. Two possibilities occur which may explain the difference in the two sets of results First, we have seen that Wright may have had two chairs in close proximity (it is not clear from his monograph whether he did), furnishing a convenient path to the supposedly barred object Second, this is the experiment, referred to earlier, in which all but two of Wright's subjects had previously taken part in an almost identical experiment in which they had nearly all taken the available object, a sort of satiation with the latter, or some other motive for change, may be the significant factor responsible for his results[5]

b Table-distance candy experiment, Stage 1. For this experiment a small room, ordinarily used as a sleeping place for children who have difficulty napping in the larger room, was employed The room opened directly into the large play-room as well as onto the hall and the children were accustomed to running in and out of it. A bridge table was so placed that when a child entered from the hall he would be opposite the midpoint of the table. The "available" candy was two inches in from the midpoint of this side of the table, while the "barred" candy was two inches in from the midpoint of the opposite side The barrier consisted only of the presence of the table, which the child had to go around in order to get the more distant piece of candy. After bringing the child in, the experimenter stood behind him and said, *"Do you see the two pieces of candy on the table? Well, you may get either one you want."*

---

[5] The same explanation appears to be applicable to the difference in results between Wright's "cage-toy experiment, stage 1" and his "cage-toy experiment, stage 2," which he attributes to another factor.

In this experiment again, the children who had participated were kept apart from those who had not

Results. Of the 20 children used in this experiment, 15 chose the available candy and 5 chose the barred one In Wright's comparable experiment, 7 out of 17 subjects chose the less accessible candy

Discussion. Here our results and Wright's are in close agreement, indicating that about a third or slightly more of the children chose the barred object under the conditions of this experiment

c. Table-distance candy experiment, Stage 2. In the experiment to which Wright gives this title, the procedure of the last experiment is duplicated except that a sieve is placed over the more distant piece of candy in order to increase the extent to which a barrier lies between it and the subject We conducted this experiment in the parlor of another nursery school. This was a room which the children go through each time they visit the head of the school, as they generally do several times a day The chairs and tables in the room were all close to the walls, and we placed a bridge table in the center As before, the children had no contact before entering the experimental room with those who had already acted as subjects

Results. All of the 20 subjects in this experiment chose the available piece of candy Nine of Wright's 11 subjects chose the barred piece

Discussion. Our results and Wright's are here sharply contradictory. No special factor is readily apparent in the procedure followed by either of us in this experiment alone which might explain the divergence of results

d. Table-distance candy experiment, Stage 3. After Wright found that a large proportion of his subjects chose the barred candy in the experiment just described, he felt that a check was needed on the possibility that the sieve was functioning as a secondary goal-object rather than as a barrier. He therefore performed a third stage of the experiment, in which an identical sieve was also placed *under* the available candy Thus, a sieve was in immediate proximity to each piece of candy but was so placed as to constitute a barrier only for the already less accessible one

We performed this experiment in the same nursery school and in the same room as the one described above in Stage 1 The conditions differed, however, in that it was not possible to isolate the children who had acted as subjects from those who had not yet done so

Results Of the 20 subjects in this experiment, 19 chose the available piece of candy and only one chose the barred piece Of Wright's 6 subjects, 5 chose the available piece and one the barred.

Discussion Our results here are in agreement with Wright's, insofar as the results from his very small number of subjects provide any basis for comparison Considering this and the previous experiment together, the most reasonable inference from Wright's results is that for his subjects the sieve was itself a goal-object, so that when a sieve was placed close only to the barred piece of candy the children went there but when each piece of candy was close to a sieve the real preference for the available piece was revealed Wright argues on the basis of "clinical impression" that this inference is not correct. If, however, we make this inference, then our results and his are perfectly consistent if we further assume, as we seemingly must from the data, that for our children the sieve did not have any appreciable value as a goal-object It remains a puzzle why for both experimenters the choice of the barred object should have been so much less frequent even than it was in the very similar experiment reported as "Stage 1," but this may be a matter of sampling error

2. *Verbal Choice Experiments*

a. Cage-toy experiment. In this experiment two toys were placed inside a large wire cage on a play-room table and two other toys judged to be equally desirable were placed outside the cage on the table, with the pairs of toys being interchanged in random order for various subjects. The cage was considered to constitute a barrier between the child and the first mentioned pair of toys Instead of being asked to get the toys himself the child was simply told, *"Tell me which one of those four toys you like best. After that, which one is the nicest,"* and so on until all four had been ranked by the child.

A very small room was used in which each child ordinarily has a brief daily examination by the doctor. The cage was made so as to conform to the specifications given by Wright for the one he used. The toys were all alike in color. Different sets of toys were used for boys and for girls. It was not possible to isolate the children after they had served as subjects.

Results. Fifteen out of 20 children gave first and second places in their rank order to the two available toys Only one child gave the first two places to the barred toys Three subjects assigned the barred toys to first and fourth places, giving the second and third places to the available ones. One subject, finally, gave the first and third places to the available toys Thus, the available toys were given the higher rating by 80 per cent of the children, the barred ones by 5 per cent, and 15 per cent of the children gave an equal rank to the two sets of toys (if the rank numbers for the two toys in each set are added or averaged) Wright reports that in his 18 subjects the barred toys were given preference by 61 per cent and the available toys were preferred by only 11 per cent, with 20 per cent giving the two sets equal ranking

b. Cage-candy experiment This experiment was analagous to the last one except that four pieces of candy were substituted for the four toys The child was told, *"Tell me which piece of candy you like best and I'll get it for you Which one do you like next best?"* and so on until he had ranked all four pieces The candy he named as his first choice was given to him after he had completed the rankings, and he was given to understand that he would receive only that one piece

This experiment was done in the parlor of a nursery school, a room which the children visited daily It was not possible to segregate the children who had already participated in the experiment.

Results Sixteen of 20 children ranked the available pieces of candy as their first and second choices. Two others gave the first place to one of the barred candies but the second and third places to the available ones One subject reversed this ordering, while the other, finally, gave second and fourth places to the barred candies Thus, the available candies were ranked above the barred ones by 85 per cent of the children, and the other 15 per cent gave equal rank to the two. With Wright's subjects in the comparable experiment, the barred pieces of candy were given higher rank by 61 per cent and available were given higher rank by only 26 per cent, while 13 per cent of his children gave equal rank to the two

c. String-suspension candy experiment This experiment was done in the same room and with the same physical arrangement as the experiment already described under an identical title among the action choice experiments, with the exception that no chair was present to mediate approach to the barred piece of candy. The experimenter said to the child, *"Tell me which piece of candy you like best and I'll get it for you."*

The subjects for this experiment were children who came to the school at a different time from those used for the action choice experiment, and attended a different group on a separate floor of the building Having no contact with the other class, they presumably knew nothing about the experiment ahead of time, and the children who had participated were kept separate from those who had not

Results. Of our 20 subjects, 17 asked for the available candy and only 3 asked for the barred one. In Wright's similar experiment, 11 children asked for the barred piece and only 5 for the available one.

d. Discussion of the verbal choice experiments. In all three of the verbal choice experiments our results differ sharply from those obtained by Wright In each case his subjects preferred the barred goal-objects while our subjects preferred the available ones In the case of the cage-toy experiment there is one difference in our procedures that might explain the difference in results Wright apparently allowed his subjects to play with the toys for a short time after they had expressed their preferences, whereas we merely asked for their preference ratings Thus, the toys must have been less of a desired goal for our subjects than for Wright's In the case of the other two experiments, however, there is no difference in our procedures that seems very likely to explain the great difference in results.

It might be reasonable to guess that Wright's subjects preferred the barred object because of prior knowledge about the situation communicated by the other children Apparently all of Wright's subjects had the opportunity to learn about the experiment from children who had already participated. It is possible that some of the first children brought in for the experiment saw the situation as involving a choice between putting forth effort and not doing so, and that they communicated this idea to the other children In this case, the rest of the children may fairly uniformly have seen the experiment as an opportunity to display effortfulness, or as a challenge to their character and to their desire to compete with the other children in doing the difficult and mature thing We had only two experiments in which the children had an opportunity to communicate with those who had already taken part and it may be that by chance none of our children talked about the experiment in this way and that some of Wright's did There is, however, no evidence that communication of this sort of idea is at all likely to have taken place among Wright's subjects.

D Summary and Conclusions

Seven experiments which had been reported by Wright on children's preference for an available object as against a barred object were repeated, although with certain modifications In general our results were contrary to Wright's in all those cases where he found a preference for the barred object and agreed with his in all those cases where he found a preference for the available object

Our results with one group of four experiments, called "action choice experiments," do not appear contradictory to Wright's if it is assumed that satiation with an available object which had previously been chosen, or that a secondary goal value of the barrier itself might cause the children to take the barred object even if they had a basic preference for the available object In the case of the other group of three experiments, called "verbal choice experiments," no such explanation offers itself as a means of reconciling the differences in our results The differences may be dependent upon cultural differences between northern and southern children, but this seems unlikely. It is more reasonable to attribute them to differences in the experimental situation and procedure which have escaped the notice of the experimenters Both experimenters obtained results which on the whole agreed with their own expectations, and it is possible that either or both of them may have unwittingly given the children cues as to which piece of candy he would like them to take, but we have no reason to believe that this happened in either group of experiments

The consistency with which our subjects in a variety of experimental situations preferred the available objects appears to argue quite conclusively against the generality of the hypothesis which Wright has put forward. At the very least our results justify the conclusion that that there is no uniform tendency to choose a barred object, and the results, taken in conjunction with Wright's, give strong initial evidence for the greater generality of an opposite tendency to choose the more available of two objects The significance of the latter conclusion is strengthened by the fact that it agrees with the general direction of results in a variety of other experiments both with animals and with human beings

Although we doubt the generality of Wright's hypothesis and feel that great inadequacies appear in some of his experimental procedures and in his interpretations of his results, we feel that his monograph has made a valuable contribution He has called attention to a form of behavior, the preference for goals difficult of attainment, which is highly important in a practical and ethical sense, and offers a challenge to any theory which seeks to explain the diversities of human behavior He has presented fairly convincing evidence that this preference appears under some circumstances in older children and in adults. We believe it would be most fruitful to plan future investigation, not on the assumption that this preference represents directly a basic principle of behavior, but on the assumption that it is quite likely to involve secondary motives acquired in the course of the child's upbringing. Looking at the problem in this way, we will sketch briefly some of the ways in which this preference might be learned by a child growing up in our society

1. There can be no doubt that preference for a barred goal-object would be found in many cases when an individual has just been enjoying an available goal and has become satiated This is the idea that seems to be expressed in the proverb, "The grass on the other side of the fence is always greener." A child may frequently have the experience of proceeding from intercourse with a satiated available object to intercourse with a barred object, and of finding that the latter is in fact at that time more rewarding, because intercourse with it is not satiated, even though intrinsically the barred object may be no more rewarding than the available one If intercourse with the barred object is not allowed to proceed to the point of satiation, the child may prefer it to the available one the next time he is brought into the situation, even though at that moment there is no satiation with the available object The preference thus established for a specific barred object might generalize to other choice situations which were sufficiently similar. For example, let us suppose that a child is usually allowed to play only in his own backyard but that he occasionally is able to climb over the fence and play for a short time in the neighboring yard. Let us suppose that the satiation with playing in his own yard is sufficient so that the neighbor's yard almost always seems more attractive to him and it is very rewarding to be able to go there to play Let us suppose then that he visits a friend's house and is taken out to the backyard to play. It is conceivable that his previous experience would generalize in such a way that he would immediately begin looking over the fence and thinking how nice it would be to go over into the next yard to play.

2. It may often happen that when a child appears to be choosing a barred object rather than an available object, the reason is solely that the barrier itself or the activity of overcoming it has a goal value Such a preference would be expected in circumstances in which delay in obtaining the eventual goal was not very frustrating, for by means of this preference the child can obtain both the eventual major goal and a subsidiary goal—

greater total reward than if he had at the outset chosen the immediately available goal-object. In this case again, the reinforcement of the tendency to choose the "barred" object may generalize to other similar situations, even if in those situations the barrier is such that the activity of overcoming it is not of itself rewarding. For example, let us imagine that a child often is given the alternative between eating a piece of candy which his parents have on hand in the house for him, or of walking down to the drug store with his father, where his father will buy him a piece of candy at the same time that he gets the evening paper The child then has the reward not only of the piece of candy but also of what we will suppose to be a highly enjoyable walk with his father. It is possible that on a future occasion, when he has the opportunity to choose between taking a piece of candy which is available at home and walking to the drug store alone to buy a piece of candy, he will choose the latter alternative because of generalization of the reinforcement previously received, even though he does not ordinarily enjoy walking by himself

3 A child often has the opportunity of discovering in a variety of particular situations that an object that is difficult to attain is better than an object that he could have obtained more easily The reinforcement received from being more rewarded for choosing an object which was barred, because it actually was more desirable than the available one, may generalize to new situations where the two rewards are in fact equivalent For example, if a hungry child is allowed to get his own afternoon snack in the kitchen he can with a minimum of difficulty take a piece of bread and eat it. It requires much more effort for him to butter the bread for himself and even more to make a sandwich. Thus, there is a considerable barrier to be surmounted before obtaining a sandwich, whereas a slice of bread is immediately available Yet the sandwich is (we will assume) more rewarding The consequent reinforcement of the tendency to choose the goal more difficult of attainment might generalize in such a way that the child would later choose, let us say, to make a sandwich for himself rather than eat some equally desirable food which was immediately available

Three hypothetical cases have been cited so far in which a learning sequence might lead the child to prefer a barred object over an available object even though the two were intrinsically equally desirable to him. In all three of these cases, however, the fact that one of the objects has a barrier in front of it is not the essential element in the situation, and the whole sequence might be rephrased without the use of the word "barrier," or any equivalent. In the first case cited, it might be said that what the child had learned was to choose a goal similar to a goal with which activity had not been previously satiated, rather than a goal similar to one with which activity had been satiated. *In the second case, it could be said that the child* had learned to engage in activity similar to activity which had previously been experienced and found rewarding. In the third case, the child could be said to have learned merely to take that route to a goal which on earlier occasions had been found to produce the greatest real reward. When these cases are phrased in such a manner, it can be seen that the mere presence of a barrier is not the essential element in the situation that would be expected to evoke the newly learned behavior. The preference which has been reinforced would be expected to generalize also to various choice situations where the presence or absence of a barrier was not an element at all. We have chosen as examples imaginary situations where the presence or absence of a barrier was still an element merely because we wished to show that preference for the barred object in such a case could be predicted from the previous experience of the subject, even though the presence of a barrier was not in itself an essential feature of the situation determining the child's reaction.

4. There is a fourth possibility now to be considered, which is that the previous experience of the child might lead to a state of affairs in which the mere fact of there being a barrier would be an important element in determining the child's preference for the barred object. There may be many occasions on which a *child receives from his social environment* an extra reward for choosing a difficult route to a goal, over and above the reward that is afforded by the primary goal itself. To take the same example that we have considered in Paragraph 3, it is likely that a child who does prepare food for himself similar to that which his mother would normally prepare for him, instead of taking the simple route of eating the food in an unprepared form, will be praised for having done so. In this particular case the mother may be directly motivated to praise the child in the interests of her own economy of effort, because if the child's behavior can be encouraged he may on subsequent occasions be willing to prepare some of his food himself which his mother would otherwise have to prepare for him. In addition, however, many parents are eager to detect and encourage any evidence of increasing maturity and independence in their children, quite apart from its direct value to the parents, simply because of recognizing that the child needs to grow up psychologically as well as physically and that the success of the

child's competition with his age mates is dependent upon his attaining a degree of maturity and independence For this reason, too, parents are likely to be proud of their children's demonstration of ability to do difficult things, and to express quite openly their pleasure Similar reactions are likely to be displayed by teachers and other persons of authority The consequent reinforcement of the tendency to perform the particular act that is thus rewarded may in time generalize to many other situations where there is a choice between a difficult and an easy route to substantially the same goal This generalization would be especially likely to occur where the situation was such as to make it closely similar to the earlier learning situations, as for example by an authoritative adult's being present and watching the child's behavior, or by the adult's using such words as "difficult," "hard," or "easy," which are likely to have been used in the parents' original praise of the child.

We believe that this fourth sequence is the one likely to be involved the most commonly in cases where a barred object is preferred over an available object. The reason is that in the first three cases it seems likely that the originally reinforced tendencies will when generalized to new situations be very frequently extinguished, and that they will consequently either die out or become highly specific In the fourth case, it seems to us that the social situation is such as to provide frequent reinforcement for the tendency in a great variety of specific situations, so that a highly generalized tendency to prefer barred objects in social situations would be expected Differential extinction of the tendency might be predicted for situations in which the child is alone, for the social reward cannot there be immediately provided But at a stage of maturity where the child has internalized social norms to a point where his own self-esteem becomes involved in the question of whether he chooses a difficult or an easy path to a goal, he may still prefer a barred object even when he is alone.

REFERENCE

1 WRIGHT, H F. The influence of barriers upon strength of motivation. *Contrib. Psychol. The.*, 1938, 1, No. 3. Pp 143

Department of Psychology and
Institute of Human Relations
Yale University
New Haven, Connecticut

Presbyterian Hospital
New York City

The Journal of General Psychology, 1944, 65, 113-131.

A SURVEY OF PSYCHOLOGICAL, SOCIAL, AND ENVIRONMENTAL DIFFERENCES BETWEEN ADVANCED AND RETARDED READERS*

Fordson Public Schools, Dearborn, Michigan

JOSEPH JACKSON

A. INTRODUCTION

1. *Foreword*

The following survey presents the results of a comparative study of advanced and retarded readers which was requested by the elementary school principals of the Fordson School System The work was carried on by a study group composed of representative reading teachers appointed by the principals of the participating schools The Bureau of Research and Informational Service served as the coordinating agency for this survey

Many psychologists, teachers, and research workers have noted that children with reading disabilities have been influenced by personality factors, emotional disturbances, and social forces which are believed dynamic enough to induce reading difficulties singly or in combination with other factors In keeping with this belief, a local survey was initiated to gain further insight into the factors inhibiting or facilitating reading ability through the study of various psychological, social, and environmental forces This study purports to determine the relation of certain factors to reading ability.

2 *The plan*

This survey is concerned mainly with the determination of psychological, social, and environmental differences that exist between the advanced readers or those found in the upper quarter (above Q_3) of each grade and the retarded readers or those found in the lower quarter (below Q_1) of the corresponding grade The aforementioned classification was determined by the city norms established by means of the Metropolitan Reading Tests in Grades 2-3 and the New Stanford Reading Tests in Grades 4-6. Since the purpose was to find these differences among the advanced and retarded readers only, the interquartile range of readers (those falling within the

---

*Received in the Editorial Office on December 7, 1942

middle 50 per cent of each grade) were omitted from this study. Moreover, only students with intelligence letter ratings of *C*— (*IQ* 90-95) and above were included in this study except seven cases of *D* letter ratings (*IQ* 83-89) found in the upper quarter in which instance these seven students were classified as advanced readers and were studied as such The reason for this last exception is that students with *D* and *E* letter ratings already comprise the majority of the personnel in the remedial classes and hence receive more than an average amount of attention This does not mean that students with *D* and *E* letter ratings were not interviewed by the teachers, because all students were interviewed who were thought to be retarded in reading regardless of letter ratings or academic success The students with *D* and *E* letter ratings were in general excluded in order to reduce differences in innate ability and to afford a better basis for the study of other factors that may influence reading ability

3 *The Procedure*

The procedure used in this study was to obtain psychological, social, and personal data concerning the advanced and retarded readers. This information was gathered through the interview technique and through an analysis of standard school records. Approximately 300 advanced readers and an equal number of retarded readers in Grades 2-6 of the Fordson School District during the second semester of the 1939-40 school year served as the subjects in this survey

At no time during the survey was the experimental method used Rather, the causal—comparative technique was employed since it was desired to establish causal relationships by comparing the circumstances associated with investigated or observed effects, and by noting the factors present in those instances in which a given effect occurred or did not occur Moreover, it is significant in some cases to begin with certain observed effects in order to explain the antecedents of these effects (1)

4 *The Treatment of Data*

The information gathered was classified under 20 headings. These 20 items are listed, explained, and discussed below.

Analyses of 16 of the items are presented in Table 1 The data were treated in a dichotomous manner Table 1 includes for 16 items the number and per cent of pupils in each category Some of the items involve a number of sub-divisions; therefore, Table 1 includes 63 percentages for

TABLE 1

NUMBER AND PER CENT OF ADVANCED AND RETARDED READERS IN THE CATEGORIES FOR EACH DIVISION OF 16 CLASSIFICATIONS OF DATA

Variables	Advanced readers		Retarded readers	
	No	%	No	%
1. Sex				
a. Girls	177	59.0	110	36.7
b Boys	123	41.0	190	63.3
2 Letter rating or *IQ* distribution				
A-IQ (118)	100	33.3	14	4.7
B (111- 17)	75	25.0	32	10.7
C+ (105- 10)	62	20.7	85	28.3
C (96-104)	47	15.7	104	34.7
C— (90- 95)	9	3.0	65	21.7
D (83- 89)	7	2.3	0	0.0
3 Language spoken at home				
a Majority of time foreign	36	14.3	46	20.6
English	216	85.7	177	79.4
b Foreign part of time	48	18.2	77	30.3
English	216	81.8	177	69.7
c. Any foreign	84	28.0	123	41.0
English	216	72.0	177	59.0
d. Majority versus part of time				
1 Majority	36	42.9	46	37.4
2 Part of time	48	57.1	77	62.6
4 Occupation of father				
a Professional (lawyer, doctor, etc.)	4	1.4	1	.4
b Semi-professional (teacher, engineer, etc)	10	3.6	3	1.1
c Specialized (accountant, salesman, etc)	27	9.7	15	5.4
d Skilled (carpenter, machinist, etc)	86	30.9	79	28.4
e Unskilled (laborer, sweeper, etc)	148	53.2	173	62.2
f Unemployed	3	1.1	7	2.5
5 Grade(s) repeated				
a Girls	5	2.8	32	29.1
b Boys	9	7.3	71	37.4
c Total	14	4.7	103	34.3
6. Grade(s) skipped				
a Girls	22	12.4	5	4.5
b. Boys	13	10.6	2	1.1
c Total	35	11.7	7	2.3
7 Personal illness				
a Childhood type (measles, whooping cough)	12	4.2	46	15.9
b Recurrent (colds, ear trouble, etc.)	3	1.0	7	2.4
c. Major type (pneumonia, diabetes, etc.)	16	5.5	12	4.2
d Unclassified	5	1.7	14	4.8
e Total	36	12.5	79	27.3

TABLE 1 (*continued*)

Variables	Advanced readers No.	Advanced readers %	Retarded readers No.	Retarded readers %
8. Family illness				
a. Parents	16	5.9	17	6.3
b. Others (sibs, grandparents)	8	3.0	4	1.5
c. Total	24	8.9	21	7.8
9. Personality classification				
a. Average	43	16.5	88	33.7
b. Good	116	44.4	54	20.7
c. Excellent	32	12.3	11	4.2
d. Sub-average (bold, untrustworthy, etc.)	23	8.8	49	18.8
e. Introverted (dreamer, reticent, melancholy, etc.)	36	13.8	43	16.5
f. High-strung (restless, nervous, flighty, etc.)	11	4.2	16	6.1
10. Semester's citizenship grades				
A	100	38.3	42	16.1
B	103	39.5	82	31.4
C+	19	7.3	21	8.0
C	31	11.9	90	34.5
C—	3	1.1	9	3.4
D	5	1.9	17	6.5
11. Physical defects				
a. No	276	95.5	244	90.7
b. Yes	13	4.5	25	9.3
12. Read to by				
a. Self	164	65.1	127	50.4
b. Others	88	34.9	125	49.6
13. Fears reported				
a. No	221	77.3	186	69.1
b. Yes	65	22.7	83	30.9
14. Family economic status				
a. Destitute	1	.6	2	1.1
b. Poor	10	5.8	23	12.6
c. Fair	77	44.5	106	58.2
d. Good	82	47.4	49	26.9
e. Excellent	3	1.7	2	1.1
15. Speech defects				
a. No	219	89.7	181	77.4
b. Yes	25	10.2	53	22.7
16. Dominant Handedness				
a. Right	215	92.3	202	95.3
b. Left	18	7.7	10	4.7

the advanced readers and an equal number for the retarded readers. The data in Table 1 were evaluated. The statistical calculations involved in the evaluation are presented in Table 2, which contains the differences between the appropriate percentages based upon the advanced readers and retarded readers. The standard errors of the differences between the percentages, the ratios of the differences to the standard errors of the differences, and the

TABLE 2

DIFFERENCES BETWEEN RETARDED AND ADVANCED READERS BASED UPON 16 CLASSIFICATIONS OF DATA

Variables compared	Differences between percentages* Adv	Ret	SE_d	$\frac{D}{SE_d}$	Chances in 100 of a true difference greater than zero
1 Sex					
Girls (minus boys)	22.3		3.98	5.60	100
2 Letter rating or *IQ* distribution					
A-IQ (118)	28.6		2.98	9.60	100
B (111- 17)	14.3		3.07	4.66	100
C+ (105-110)		7.6	3.50	2.17	98
C (96-104)		19.0	3.46	5.49	100
C— (90- 95)		18.7	2.58	7.25	100
D (83- 89)	2.3		.87	2.64	99
3 Language spoken at home					
a Majority of time foreign		6.3	3.49	1.81	96
b Foreign part of time		12.1	3.74	3.24	100
c Any foreign		13.0	3.84	3.39	100
d Majority versus part time	5.5		6.94	.79	77
4 Occupation of father					
a Professional	1.0		.80	1.25	89
b. Semi-professional	2.5		1.28	1.95	97
c Specialized	4.3		2.23	1.93	97
d Skilled	2.5		3.87	.66	74
e Unskilled		9.0	4.17	2.16	98
f. Unemployed		1.4	1.13	1.24	88
5 Grade(s) repeated					
a. Girls		26.3	4.50	5.84	100
b Boys		30.1	4.22	7.13	100
c. Total		29.6	3.00	9.87	100
6 Grade(s) skipped					
a Girls	7.9		3.17	2.49	99
b. Boys	9.5		2.88	3.30	100
c Total	9.4		2.05	4.59	100
7 Personal illness					
a Childhood type		11.7	2.45	4.78	100
b Recurrent		1.4	1.07	1.31	90
c Major	1.3		1.79	.73	76
d. Unclassified		3.1	1.47	2.11	98
e Total		14.8	3.26	4.54	100
8. Family illness					
a Parents		.4	2.06	.19	58
b Others	1.5		1.27	1.18	87
c. Total	1.1		2.38	.46	67
9 Personality classification					
a. Average		17.2	3.72	4.62	100
b Good	23.7		3.97	5.97	100
c. Excellent	8.1		2.38	3.40	100
d Sub-average		10.0	2.99	3.34	100
e Introverted		2.7	3.14	.86	80
f High-strung		1.9	1.93	.98	83

*Percentages identify group having majority difference.

TABLE 2 (*continued*)

Variables compared	Differences between percentages* Adv	Ret	SE_d	$\frac{D}{SE_d}$	Chances in 100 of a true difference greater than zero
10. Semester's citizenship grades					
A	22.2		3.77	5.89	100
B	8.1		4.17	1.94	97
C+		.7	2.33	.30	62
C		22.6	3.56	6.35	100
C—		2.3	1.29	1.78	96
D		4.6	1.74	2.64	99
11. Physical defects		4.8	2.15	2.23	99
12. Read to by others		14.7	4.35	3.38	100
13. Fears reported		8.2	3.75	2.19	99
14. Family economic status					
a Destitute		.5	.80	.63	73
b Poor		6.8	3.03	2.24	99
c Fair		13.7	5.26	2.60	99
d Good	20.5		5.02	4.08	100
e Excellent	.6		1.25	.48	69
15. Speech defects		12.5	3.36	3.66	100
16. Dominant handedness—right		3.0	2.30	1.30	90

*Percentages identify group having majority difference.

chances of true differences greater than zero are also portrayed in Table 2.

Statistical analyses of three of the items are given in Table 3. These three items were separated from the first 16 because the treatment of the data is not the same. For each item in Table 3, the means, medians, standard deviations, and standard errors of the means and medians were computed. These calculations were made separately for advanced and retarded readers. Then the two groups were compared on each item of information on the basis of differences between the appropriate means and medians. Table 3 includes these differences as well as the standard errors of the differences, the ratios of the differences to the standard errors of the differences, and the chances of true differences greater than zero. Both measures of central tendency, the mean and the median, were calculated for each of these items because the two measures yield a better comparative analysis than either one alone would. This is particularly true for these three items because they represent types of data which are definitely skewed.

The last item of information which was considered in this study pertains to the magazine preferences and reading of the advanced and retarded readers. These data are summarized in Table 4. The magazines read by

TABLE 3

DIFFERENCES IN FAMILY SIZE, PARENTAL SCHOOL GRADE ATTAINMENT AND DAYS OF ABSENCE BETWEEN ADVANCED AND RETARDED READERS

Variables	Advanced readers Central tendency	SD	SE	Retarded readers Central tendency	SD	SE	Diff	SE_d	$\frac{D}{SE_d}$	Chances in 100 of a difference greater than zero
17 Family size										
Mean	5.27	1.70	10	5.45	2.00	12	18	16	1.13	87
Median	4.87	1.70	13	6.00	2.00	.15	1.13	.20	5.65	100
18 Parental attainment in school										
Fathers—										
Mean grade	9.50	2.88	55	7.08	3.99	45	2.42	57	4.25	100
Median	9.50	2.88	44	8.20	3.99	.56	1.30	71	1.83	97
Mothers—										
Mean grade	9.34	3.50	39	8.00	3.40	.39	1.34	55	2.44	99
Median	8.00	3.50	.49	8.56	3.40	49	56	69	81	79
19. Days absent										
Mean	6.02	6.77	44	6.08	6.53	40	51	59	10	54
Median	3.69	6.77	.55	4.26	6.53	50	57	74	77	78

TABLE 4

PERIODICALS PREFERRED AND READ BY THE RETARDED AND ADVANCED READERS IN GRADES 4 THROUGH 6 IN ORDER OF PREFERENCE

Periodicals	Frequency of usage Reader Ret	Adv.	Rank order Reader Ret	Adv	Difference in rank order Ret	Adv
Life	31	19	1.0	2.5	—	1.5
True Story	21	10	2.0	9.5	—	7.5
Comic Magazine	18	10	3.0	9.5	—	6.5
Liberty	15	16	4.0	4.0	0	0
Child Life	10	23	6.0	1.0	5.0	—
Collier's	10	19	6.0	2.5	4.5	—
Movie Magazines	10	11	6.0	6.0	0	0
Look	9	4	8.0	20.0	—	12.0
Boy's Life	8	10	9.5	9.5	0	0
Popular Mechanics	8	12	9.5	5.0	4.5	—
Daily Paper	7	5	11.5	17.5	—	6.0
Saturday Evening Post	7	9	11.5	13.0	—	2.5
Good Housekeeping	5	6	14.0	15.5	—	1.5
Ladies' Home Journal	5	10	14.0	9.5	4.5	—
American	5	3	14.0	24.0	—	10.0
Popular Science	4	3	17.5	24.0	—	6.5
Better Homes and Gardens	4	4	17.5	20.0	—	2.5
American Boy	4	—	17.5	—	—	—
Mystery	4	—	17.5	—	—	—
National Geographic	3	1	20.5	40.5	—	20.0
Motor News	3	—	20.5	—	—	—
Grade Teacher	2	—	24.5	—	—	—
Children's Activities	2	3	24.5	24.0	5.0	—
Popular Aviation	2	2	24.5	29.5	—	5.0
*Wee Wisdom**	2	10	24.5	9.5	15.0	—
Detective	2	4	24.5	20.0	4.5	—
Radio Story	2	—	24.5	—	—	—
Other (18)**	1	—	36.5	—	—	—
W Home Companion	—	10	—	9.5	—	—
American Girl	—	7	—	14.0	—	—
McCall's	—	6	—	15.5	—	—
Reader's Digest	—	5	—	17.5	—	—
Playmate, Parents	—	3	—	24.0	—	—
Click, Fishing and Hunting, Open Road for Boys, Sports, Red Book	—	2	—	29.5	—	—
Other (20)***	—	1	—	42.5	—	—

*Not a periodical, included because of rank.

**Eighteen magazines had a frequency of one.

***Twenty magazines had a frequency of one.

the two groups are listed The number of pupils in each group who read each magazine is indicated These numbers were then used to obtain the rank order of each magazine on the basis of the frequency with which they were read by these two groups of pupils. The differences between the rank orders of the advanced and retarded readers were also obtained and included in Table 4

B. RESULTS OF A COMPARISON OF ADVANCED AND RETARDED READERS UPON 20 CLASSIFICATIONS OF DATA

The statistical results involving the 20 groupings of data used in this study are presented in Tables 1-4 Advanced and retarded readers were compared in terms of percentages on the first 16 items. These results are included in Tables 1 and 2 Each of these 16 items is discussed below in the same order that they are listed in Tables 1 and 2

1 *Sex*

The first comparison between the two groups pertains to sex or to the number of boys and girls Of the 300 advanced readers surveyed, 59 per cent are girls and 41 per cent are boys The retarded readers consist of 36 7 per cent girls and 63 3 per cent boys The data indicate that the incidence of boys among the poor readers is too great to be explained on the basis of chance alone The data reveal that there are approximately six boys to four girls in the retarded group; however, in the advanced group this ratio is reversed A glance at the data in Tables 1 and 2 shows that the percentage of boys in the retarded group exceeds the percentage of boys in the advanced group 22.3, or one may say that the percentage of girls in the advanced group exceeds the percentage of girls in the retarded group 22.3 The standard error of this difference of 22.3 is 3 98, and the difference is 5 60 times the standard error of the difference A ratio of three is generally accepted as an indication of a true difference greater than zero; therefore, the data reveal a statistically significant difference between the sexes in relation to reading ability The present study does not demonstrate that sex has a causal effect upon reading ability It is possible that some of the other factors included in this study may account for all or a large portion of the differences in reading ability of the two groups under consideration The data presented here suggest that sex should be one factor under consideration in an investigation of the causes of reading ability or disability

2 *Letter rating or IQ distribution*

The second comparison between the two groups involves the *IQ*'s which were obtained from the administration of the Detroit Intelligence Tests, Forms *C* and *M* The *IQ*'s were grouped into letter ratings as follows· *A*, 118 and above, *B*, 111-117, *C*+, 105-110, *C*, 96-104, *C*—, 90-95, *D*, 83-89, and *E*, 82 and below The percentage of advanced and retarded readers with the various letter ratings was calculated.

The advanced group of readers is dominated by a large proportion of *A* and *B* letter ratings. The differences between the percentages of these ratings in the two groups are statistically significant This fact is borne out by the critical ratios of 9.60 and 4 66 respectively for these letter ratings The retarded readers on the other hand have a significantly larger proportion of *C* and *C*— letter ratings. The critical ratios involving these ratings are 5.49 and 7 25 The data tend to show that while the more intelligent students constitute a large proportion of the upper quarter of readers, the average students occupy the lower quarter of readers with a greater than expected rate of incidence since normally this group would be expected to achieve a reading level within the interquartile range of readers Of six comparisons between advanced and retarded readers, four reveal differences that are statistically significant The remaining two differences are nearly statistically significant Intelligence seems to be a very important factor in relation to reading ability It may be possible that many of the differences between the advanced and retarded readers are a function of differences in intelligence This possibility cannot be demonstrated by a causal-comparative approach It must be attacked by a series of studies using the experimental method It is quite noteworthy that 43 7 per cent of the retarded readers have *IQ*'s of 105 and above This fact suggests that intelligence is not the only factor causing differences in reading ability, therefore, one must search for additional factors or combinations of factors of a causal nature

3 *Language Spoken at Home*

A comparison of the language or languages spoken in the homes of advanced and retarded readers brings out significant critical ratios in two categories. The first notable difference is for the *part-time* foreign speaking group of retarded readers which produces a ratio of 3 24 while the second difference is for the combined *majority* and *part-time* foreign speaking group which yields a significant ratio of 3.39. For the purposes of this study, majority means more than half of the time, while *part-time* means half of the time or less. The home-room teachers who conducted the interviews classified the responses pertaining to language spoken at home. These results would perhaps indicate that the retarded readers are not exposed to an English vocabulary in the home to the same degree as are advanced readers and, hence, would not possess the same opportunity for vocabulary, word meaning, and word recognition development as do advanced readers. Consequently, whenever the retarded readers are measured on the above three phases of reading, they are in a sense handicapped

4 *Occupations of Parents*

In general, the comparisons of the two groups on the basis of occupational status of fathers reveal slight differences between percentages In both groups the children are mainly from homes of skilled and unskilled workers It is of special interest to note that the differences involving the skilled and unskilled groups yield low critical ratios The social and economic status of the home as roughly measured by paternal occupation seems to bear little relation, if any, to differences in the reading ability of the two groups

5 *Grade(s) Repeated*

That the retarded readers repeat one of the first six grades more often than is true of the advanced readers is not due to chance factors. The critical ratios of 5.84 for girls and 7 13 for boys demonstrate that the differences are statistically significant. The 71 boys and the 32 girls represent 37.4 and 29 1 per cent, respectively, of the pupils in the retarded group who have repeated one or more grades; on the other hand, the advanced readers repeat one of the first six grades at a very small rate This is emphatically borne out by the high critical ratio of 9 87 for the group as a whole.

The critical ratios of 5 84 and 7 13 involving girls and boys, respectively, take on more significance when it is found that 16, or 50 per cent, of the 32 girls and 55, or 77 per cent, of the 71 boys have a letter rating of *C* (*IQ* 96-104) or higher. A breakdown of the grades during which most repetition of schooling occurs reveals it to be the period beginning with Grade 1*B* and ending with the first half of Grade 3 or Grade 3*B*.

The above facts are consistent with the analysis of readers by sex. The severity of the problem signifies that more attention should be given to it by the regular as well as the experimental schools Certainly one aspect of this problems begins with an inventory of each student as employed in this study.

6 *Grade(s) Skipped*

In terms of the incidence of skipping a grade, the girls surpass the boys in the advanced reader group 22 to 13 or 12.4 per cent and 10.6 per cent of each sex, respectively.

The difference involving the boys is significant as indicated by a critical ratio of 3 3, the difference being in favor of the advanced reader group, which result is also in agreement with the total critical ratio of 4.59. This suggests that reading is a main criterion of promotion in the first three if not in the first six years of school attendance, and since boys form the majority of retarded readers they likewise form the majority of the repeaters.

7 *Personal Illness*

A comparison of the health history of the cases reveals a larger frequency of illness among retarded readers in the more common types of children's diseases such as measles, whooping cough, diphtheria and the like. The critical ratio involving these respective illnesses is 4 78 which demonstrates that the difference is statistically significant Moreover, the group difference indicates that the retarded readers have a greater incidence of illness than do advanced readers which fact is embodied in the significant ratio of 4 54. The data indicate that illness in itself or in combination with other factors may promote problems of reading retardation. For the recurrent and major types of illness no significant group differences were found

8 *Family Illness*

A study of the family health status among both groups of readers reveals an insignificant amount of difference; the critical ratios were 19 for parents and 1 18 for others as sibs and grandparents.

9 *Personality Classification*

On the basis of interviews, the home-room teachers wrote brief summaries of the personality traits of the pupils. The writer classified each pupil into one of six categories on the basis of the material obtained from the interviews made by the teachers. A comparison of the personalities of advanced and retarded readers yields several significant results In the "average" personality classification the critical ratio of 4 62 tends to indicate that the retarded reader attains that rank with greater regularity than does the advanced reader Furthermore, this personality level is in the majority of cases the ultimate attainment of the retarded readers On the other hand, the advanced readers as revealed by the critical ratios of 5.97 and 3 40 dominate the "good" and "excellent" ratings on personality which indicate that the advanced reader is not only ideationally superior but also superior in personality traits as evaluated by classroom teachers

A study of the negative aspects of the personality classifications as the "sub-average" (bold, untrustworthy); the "introverted" (reticent, melancholic, dreaming), and the "high-strung" (restless, nervous, flighty); shows differences on the side of retarded readers although only the sub-average group yields a statistically significant difference as evidenced by a critical ratio of 3 34. This fact shows that among retarded readers one may expect a larger than chance number of individuals who possess personality traits that deviate considerably below average.

10 *Citizenship Grades*

An analysis of the citizenship grades of advanced and retarded readers disclose group differences paralleling those of the personality classification The advanced readers receive a greater proportion of "*A*" citizenship grades. The difference between the advanced and retarded readers is statistically significant as demonstrated by the critical ratio of 5 89. On the other hand, the retarded readers attain average semester grades in citizenship in much greater number The critical ratio is 6 35 involving the *C* grade Only 55.5 per cent of the retarded readers have attained a citizenship grade above the average (*C*+ and higher) as compared to 85 1 per cent of the advanced readers As a conclusion it may be said that the advanced readers attain better citizenship grades in the classroom than do the retarded readers This fact would indicate that school as well as personal adjustment on behalf of retarded readers is inferior to that of the advanced.

11. *Physical Defects*

Slight differences were found in the frequency of physical defects current among retarded and advanced readers Among the 269 retarded readers for whom data were available, 25 cases of physical defects occurred while in the 286 advanced readers surveyed 13 cases of physical defects were found

In this study physical defects include not only sensory ones such as visual and auditory defects, but also bodily deformities and marked cases of over-or-under-development These defects range from the less serious to a number of severe types The critical ratio of 2 23 shows the group differences to be insignificant.

The most noticeable difference between defects common to both groups of readers is that related to body structure Of the 25 defects listed for retarded readers, 11 are concerned with body size such as over-or-undersize, or underweight. This information is important for analyzing adequately various causes of reading disability since it may suggest that the teacher be continually informed regarding the student's need for medical analysis and follow-up.

12. *Read to by Self or Others*

Obtained data for the 252 retarded and advanced readers surveyed shows that 127 retarded and 164 advanced readers did not have others read to them. The difference yields a critical ratio of 3 38 which indicates statistical significance. This difference means that the retarded reader is less active in reading than the advanced reader. No attempt was made to quantify the amount of reading

13. *Fears Reported*

The home-room teachers questioned each pupil during the interviews to determine the incidence of fears. The 83 retarded readers reporting one or more fears comprise 30.9 per cent of the 269 persons surveyed in that group as compared to the 65 advanced readers who comprise 22.7 per cent of the 286 persons surveyed in the advanced reader group, thus indicating that fears are somewhat more prevalent among retarded than among advanced readers. The critical ratio of 2 19, however, indicates that this group difference is not significant. Perhaps of more interest are the nature and types of fears listed For example, the majority are the vague, indefinite type (shadows, guns, dark, animals, etc.) justifying neither a physical nor logical basis for their existence They show, moreover, an early conditioning of the individual through erroneous conceptions of environment or experience without the individual's knowledge of the ultimate effect that such misconceptions may cause to one's mental associations

The need for the above information becomes apparent from the types of fears presented by the retarded as well as the advanced readers. While the relation of fears to reading ability or disability does not seem evident, it nevertheless appears in specific cases so as to justify some consideration.

14. *Family Economic Status*

An analysis of the family economic status of advanced and retarded readers reveals a statistically significant difference in favor of the advanced reader in the "good" classification in which instance the critical ratio is 4 08 The differences among the remaining categories are small but nevertheless are on the side of the advanced readers These data suggest that reading ability and economic status are related and the differences common to both groups of readers may be conditioned by such factors as family background, education of parents, luxuries and the like. However, to ascertain more adequately the effect of the economic factor upon reading ability, various specialized types of techniques must be utilized.

15. *Speech Defects*

Speech defects were found to be more frequent among retarded than among advanced readers The critical ratio of 3.66 designates this frequency to be above the expected rate due to chance Since the tests administered measured the rate and comprehension of silent reading and not oral ability, it is distinctly noteworthy that reading retardation may occur on the basis of a speech defect.

16. *Dominant Handedness*

The home-room teachers classified the pupils according to dominant handedness on the basis of their classroom observations and interviews A study of dominant handedness fails to disclose any factors operating in reading achievement The majority of retarded and advanced readers are of right-hand dominance. The advanced readers were found to outnumber the retarded readers in left-hand dominance by three per cent, which, if anything, would disprove any theory that the poor readers are of left-hand dominance.

17. *Family Size*

Due to the skewness of the distribution of family size, the median is preferred as the measure of central tendency to denote the typical situation. The typical family of retarded readers is represented by 6.0 individuals while the typical family of advanced readers is represented by 4 87 individuals The median difference of 1.13 is statistically significant The critical ratio is 5.65 (the critical ratio involving the means is 1.13).

18 *Parental Attainment in School*

The amount of education of the fathers of advanced readers surpasses that of the fathers of retarded readers at all levels of school attendance For example, 70 fathers of retarded readers terminated their education in one of the elementary grades (1-6) while 29 fathers of advanced readers completed their education in one of the same grades This difference naturally reversed the results in the comparison of the secondary school grade (7-12) attainment as 101 fathers of advanced readers attended the junior or senior high school as compared to 72 fathers of retarded readers Moreover, it was found that 49 fathers of advanced readers attended high school at some time as contrasted to 33 fathers of retarded readers Neither of these figures could be incorporated into the central tendency since grade attainments were not specified Furthermore, 15 fathers of retarded readers had received no formal schooling as compared to only one parent of an advanced reader.

The mean grade attainment for the fathers of advanced readers is 9 5 and for those of the retarded readers is 7.08 The critical ratio involving the difference between means is 4 25.

The education of the mothers of advanced readers is also somewhat superior to the education of the mothers of retarded readers. The mean grade completed by the former is 9.34; by the latter, 8 0. The difference is not significant; the critical ratio is 2 44 Here again 42 mothers of advanced readers completed one of the secondary school grades (7-12) as compared to 29 mothers of retarded readers These data could not be incorporated into the calculation of the means

Several important factors are embodied in the analysis First, it may be said that the advanced reader is influenced to a degree by better education of the parents which would be reflected in educational, cultural, and literary interests of the family Second, the extent of parental education bears some relationship to family economic and occupational status which tends to raise the standards and stabilize the home of the advanced readers.

19 *Days Absent*

A study of school attendance during the second semester of the 1939-40 school year fails to reveal any significant group differences. Retarded readers are absent slightly more than are advanced readers This, of course, is consistent with other comparisons in the survey.

Since individuals in both groups were found with an absence of 10 days or more, it seems problematical whether reading retardation can occur on the basis of absence alone except perhaps in the more extreme cases. The low critical ratios seem to bear this out

20 *Periodicals Preferred*

Various differences were found in the preferences for periodicals among the retarded and advanced readers The most notable distinction is the preference of retarded readers for pictorial or illustrated magazines such as *Life, True Story, Comic Magazine, Look,* and *National Geographic* which rank 1, 2, 3, 8, and 20.5 in order of preference, but rank 2 5, 9 5, 9 5, 20 and 40 5 in order of preference among advanced readers This shows that the retarded reader is interested in illustrated material rather than the purely reading type. On the whole, while the retarded and advanced readers read practically the same periodicals, they differ only in the rank order of preference Many of these periodicals are definitely of the adult type

Other important aspects of this phase of the survey revealed that the advanced readers are more familiar with periodicals than are the retarded readers for they mentioned two or more periodicals more often than did retarded readers Moreover, the non-readers of peiodicals in Grades 4-6 among the advanced readers are approximately 1 0 per cent as contrasted to 21.1 per cent among the retarded readers It was also found that the home periodicals exert a stronger influence upon the students than do school periodicals. Of the 257 choices of periodicals listed by advanced readers in Grades 4-6, 75, or 29 1 per cent, indicate magazines found on the typical, local, elementary library list, while among the 221 choices of retarded readers are listed 49, or 22 per cent, of those common to this same list This fact may indicate that the advanced reader is more thorough in using the library. The data suggest that advanced readers select materials that are suitable for them in content as well as in difficulty Consequently, the advanced readers are able to profit more by their magazine reading

The problems involved in the leisure time reading of students are varied and complex and certainly not limited to the influence of school guidance alone Closely allied are the factors of interest, family economic status, parental school attainment, adolescent siblings, and the like which invariably affect those who are younger and inexperienced. The *types* of magazines read by the retarded and advanced readers alike suggest that this influence is active even at the fourth grade level and perhaps below

C CONCLUSIONS AND RECOMMENDATIONS

1 Various psychological, social, and environmental factors operate constantly throughout the school attendance of the students It is important

that these factors be known and understood by all the teachers concerned.

2 This survey reveals the necessity for greater attention to the extraneous factors influencing reading success such as personality traits, home conditions, etc.

3. This information should not be considered as primarily of interest to teachers of reading but should be acknowledged as a part of the general educational process. Therefore, the information should be included as a part of the student's cumulative record necessary in all succeeding grades

4. This survey indicates a definite need for the development of a counseling technique adapted to the age and grade level of all groups. A practical counseling system would begin with an individual survey such as the one used in this study.

5. A formal remedial program in reading should begin at a primary level since a large number of failures were found to occur within Grades 1*B*-3*B*

6. Because of the maturity differences of the two sexes, male readers still comprise the majority of the retarded or poorer readers and also a greater number of school failures. This reveals the need for further study of the sex factor for reading success.

7. Reading disability is not necessarily limited to students with low letter ratings since 131 of the 300 retarded readers surveyed had letter ratings of *C*+ or higher (*IQ* 105 and above). Such cases should be isolated and diagnosed early in their school career.

8 Of equal importance are the cases of personality defects common to both groups of readers who are in constant need of special guidance Co-operation with the home is highly important especially in the more serious cases

9. A survey of the type of periodicals preferred by both groups of readers, as well as the relatively high percentage of non-readers discovered among retarded readers surveyed, reveals that greater emphasis should be placed upon this aspect of reading. Moreover, the correlation of materials in accepted periodicals, newspapers, and the like should become at least an occasional part of the reading procedure.

10. It is found that a proportion of both advanced and retarded readers are dependent upon the reading of others It is important for the teacher to determine the extent to which pupils are exposed to this undirected oral reading

11 The survey seems to indicate that fears, worries, failures, introversions and the like exist to a degree which merits attention in connection with

reading difficulties. Of special importance is the knowledge on behalf of the teachers of student fears or worries arising from educational situations as failures, aversions, mispronunciations, and the like.

12 The retarded readers manifest less interest in reading than do the advanced reader. It is necessary to know the factors that condition this lack of interest such as dislikes, orientation, inertia, disability, hobbies, etc One should fashion reading assistance to fit the cause

13. The causes of reading ability or disability are many and intertwined rather than individual and isolated Therefore, continual study is needed.

14. One should consider a reading survey questionnaire of the type used in this study not as a new entity in the field of education, but rather as a part of an expanding educational program.

REFERENCE

1 GOOD, C V., BARR, A S, & SCATES, D E. The Methodology of Educational Research New York D. Appleton-Century, 1936 Pp 882.

5104 Porath Street
Dearborn, Michigan

1

The Journal of Genetic Psychology, 1944, 65, 133-145

BEHAVIOR EFFICIENCY AS A DETERMINING FACTOR IN THE SOCIAL ADJUSTMENT OF MENTALLY RETARDED YOUNG MEN*[1, 2]

Wayne County Training School, Northville, Michigan

SIDNEY W. BIJOU

Studies concerned with the empirical meaning of test interrelationships have yielded several valuable clinical concepts, one of which is behavior efficiency. This construct pertains to an area of behavior described by such overlapping terms as mental stability, mental control, mental organization (or disorganization), and mental dysfunctioning. Specifically, behavior efficiency is defined by the discrepancy between two selected tests measuring behavior under contrasting conditions (3). In clinical practice, one test has usually been a standardized performance scale such as the Army Performance Scale (12) or the Grace Arthur (4, 7), the other, a verbal power test[3] such as the vocabulary sections of the 1916 or 1937 revisions of the Stanford-Binet When the disparity is in favor of the performance rating, behavior efficiency is said to be high, when it is in favor of the vocabulary test, behavior efficiency is low. Neither type of test mentioned is entirely satisfactory for defining behavior efficiency, but their use has been forced inasmuch as they are the best available instruments for this purpose. The concept is adequately defined by the relationship between a performance test which is scored solely on speed, accuracy, and relevancy, and a vocabulary test which is evaluated on qualitative and quantitative considerations Until a scale is constructed for the specific purpose of measuring behavior efficiency, the regular clinical tests mentioned above must suffice

Individuals who are extremely disorganized,—psychotics, for example, are found to be inferior in behavior efficiency. In an investigation by Jastak (12) 84 state hospital patients with various psychiatric diagnoses had the follow-

---

*Received in the Editorial Office on December 11, 1942.

[1] From the Wayne County Training School, Northville, Michigan, Robert H Haskell, M D, Medical Superintendent, Department of Research, Thorleif G Hegge, Ph D, Director of Research and Education

[2] A condensation of this paper was presented at the Seventeenth Annual Meeting of the Midwestern Psychological Association, St Louis, Missouri, May 1942

[3] Power tests are tests which gauge the "degree of an individual's ability in a given field, measured by the difficulty of tasks which he can perform without regard to time" (17)

ing psychometric pattern: Terman Vocabulary Quotient 97, Stanford-Binet Quotient 85, and Army Performance Quotient 73. That this group rates low in behavior efficiency is indicated by the disparity of 24 points in favor of the Terman Vocabulary Test. In studies involving state hospital subjects, the Stanford-Binet characteristically falls in an intermediate position and is included to obtain a more comprehensive psychological analysis, i.e., some indication of functional intelligence as well as behavior efficiency. Prison recidivists, prototypes of a group having histories of serious social conflicts, have also been found to manifest inferior behavior efficiency (1). Data such as these strongly suggest that poor behavior efficiency is associated with a low level of personal-social adjustment.

Heretofore, measures of behavior efficiency have been related mainly to the behavior of extreme clinical types. The concept, to be useful in general psychological analysis, must be shown to be a personality characteristic existing in varying degrees in all individuals and related to a wide range of personal-social adjustment levels. This demand for pervasiveness stems from the current belief that personality traits of normal and abnormal individuals differ quantitatively, not qualitatively (6).

If it is reasonable to suppose that behavior efficiency is a trait common to all individuals and that it is related to level of personal-social adjustment, then given a group relatively homogeneous in age, intelligence, and socioeconomic background and varying in adequacy of social adjustment, those showing highest behavior efficiency should be among those making the best social adjustment, while those with low behavior efficiency should be the poorest adjusted. Because of the nature of the Wayne County Training School[4] and the thorough psychological and social services rendered, it was possible to test the above hypothesis on a group of subjects who had once been in residence here. This, then, is the primary purpose of the present investigation. The psychometric data compiled were also suitable for evaluating other tests, alone and in patterns, as they relate to social adjustment. This is the secondary purpose of the study.

The subjects consisted of 136 high grade mentally retarded boys who were paroled from the Training School between July, 1927, and July,

---

[4] The Wayne County Training School admits only high grade mentally retarded children. A majority of those accepted display problem behavior which may be described as pre-delinquent. "The prime purpose of the Training School is to ameliorate the disturbing social characteristics of selected higher grade feebleminded and mentally retarded children sufficiently if possible to permit them to return to the community" (9).

1933, (13). Personal contact with these boys was made by social workers between four and eight years after parole for the purpose of evaluating social adjustment. The salient facts concerning the group may briefly be described as follows:

Mean age on admission to the School—13 years 7 months.
Mean age at time of initial performance test—14 years 7 months.
Mean Binet *IQ* while in residence—67.
Mean age at time of parole—16 years 11 months.
Mean period of time in society—6 years 5 months.
Mean age at time of social survey—23 years 4 months.

The social workers' accounts, which followed a standard topical form, inventoried number and description of jobs held, money earned and money saved, physical appearance of the home, family, and marital harmony, leisure time activities, number of arrests and convictions, etc. These summaries, covering from one to four typewritten pages, were read independently by two judges for the purpose of rating each case on a five-point social adjustment scale. Although this instrument was created expressly for the investigation, it was not designed to take into account the mental limitations of the group. In other words, the social adjustment of these mentally retarded boys was judged on the same basis as a group of average boys would be. As the following description of categories indicates, ratings depended solely upon what the subject meant to society in terms of his demands and contributions.

Rating of 1. A burden to society: commits one or more major breaches of the law or is involved in a sequence of minor illegal acts; is continually enmeshed in social entanglements; is for the most part dependent on others for his economic and social needs.

Rating of 2. A partial burden to society: is involved in social difficulties (legal and conventional); requires frequent aid and supervision; works sporadically or frequently but earns too little to support himself.

Rating of 3. Almost socially sufficient: may have an occasional minor difficulty (social or legal); may need occasional aid from members of his family or from social agencies; works most of the time.

Rating of 4. Maintains himself and his immediate dependents· adequately takes care of the economic needs of dependents and himself; is capable of conducting himself in a socially acceptable manner; has no conflicts with social or legal forces.

Rating of 5. A social contributor· not only maintains dependents and himself but donates his time, services, or funds to other indivdiuals or groups; possesses savings; owns sizable pieces of property such as a car, a plot of land, or a home.

Since there were only five scale values to bridge a tremendous range of social behavior, the step intervals were of necessity very wide, and differences between points were easily discernible, thus enabling the judges to arrive at the same decision in 94 per cent of the cases. Summaries assigned different ratings were reconsidered jointly to reevaluate the crucial aspects of the case. If this procedure failed to culminate in an agreement, the case was discarded. Cases were also eliminated when a scale value could not be assigned with a high degree of confidence because of special circumstances or insufficient data. Because of these considerations and because some cases did not have complete psychological records, the investigation began with 202 cases and ended with 136. Of this number, 23 were assigned a social adjustment rating of 1; 41 a rating of 2; 29 a rating of 3; 37 a rating of 4; and 6 a rating of 5.

To learn whether age or length of time in the community were important factors determining the social adjustment ratings, the following group means were calculated (Table 1):

TABLE 1

Group	1	2	3	4	5
CA at parole	17-2	16-8	17-2	16-10	17-1
Time in society	6-0	6-4	6-5	6-8	6-6
Age at time of social study	23-2	23-0	23-7	23-6	23-7

It is apparent from these figures that neither of the above-mentioned factors influence category placements.

Records of psychological tests administered while the subjects were at the Training School supplied the material for measuring behavior efficiency and for other psychometric evaluations. It will be recalled that behavior efficiency is measured by the disparity between a speed-accuracy-relevancy performance test and a verbal power test. The speed-accuracy-relevancy test, which we shall refer to as the *SAR* Performance test, was derived from the first Pintner-Patterson Scale (long form) given after admission. Seven performance subtests were selected on the basis of three criteria; namely, that the item was long enough to be reliable, that it was discriminative over a wide range, and that it was scored on some combination of time, moves, and errors. Fulfilling these requisites were the following: Mare and Foal, Five-Figure Formboard, Two-Figure Formboard, Casuist Formboard, Knox Cubes, Sequin Formboard, and the Substitution Test. The seven tests yielded 11 mental ages for calculating a median mental age as described by Pintner and Patterson (16). *SAR* performance quotients were obtained in the usual manner, using 16 years as maximum growth—the procedure adhered to at the time these tests were given.

The second test required for measuring behavior efficiency—the verbal power test—was the Vocabulary section of the 1916 revision of the Stanford-Binet. Raw scores on the Terman Vocabulary Tests were taken from the Stanford-Binet administered nearest in date to the initial Pintner-Patterson Performance Scale, and by a method of interpolation (11), Vocabulary ages and quotients were derived.

The subjects were grouped into the five categories depending on the social adjustment rating, and mean *SAR PQ*'s and Terman Vocabulary *Q*'s were calculated for the Stanford-Binet, Pintner-Patterson, Stanford Achievement, and Stanford reading and arithmetic attainment, considered separately. Social adjustment ratings were then related to measure of behavior efficiency, and the other test values were viewed separately and in patterns.

RESULTS

Table 2 and Figure 1 show the mean *SAR* Performance *Q*'s and the Terman Vocabulary *Q*'s for each of the five adjustment groups. Also included in Table 2 are the mean differences between these sets of scores—the measures of behavior efficiency. It is apparent that the *SAR Q*'s and behavior efficiency ratings vary directly with social adjusment ratings, while the Vocabulary scores remain relatively constant, i.e., the differences between group Vocabulary scores are statistically unreliable. Tests of significance, based on small sample theory (15), were applied to the inter group differences in behavior efficiency, e.g., the behavior efficiency rating for Group 1 versus the behavior efficiency rating for Group 2; the behavior efficiency rating for Group 1 versus the behavior efficiency rating for Group 3, etc. All but one (the difference between Groups 3 and 5) of the differences be-

TABLE 2
MEAN GROUP SCORES FOR *SAR* PERFORMANCE, TERMAN VOCABULARY, AND BEHAVIOR EFFICIENCY

Mean test scores	Adjustment groups 1	2	3	4	5
SAR Performance *Q*'s	70.61	77.00	84.97	87.30	93.17
Terman Vocabulary *Q's*	62.13	66 03	67.52	65.74	65.17
Behavior efficiency (*SAR Q—TVQ*)	8.48	10.97	17.45	21.56	28.00

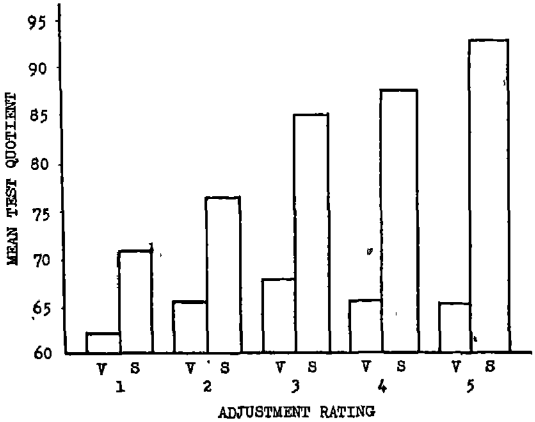


FIGURE 1

Mean Terman Vocabulary (*V*) and *SAR* Performance (*S*) Scores for Each Social Adjustment Category

tween groups, one or more categories apart, are statistically reliable at the three per cent level of confidence. Table 3 shows the values of the differences for the non-adjacent categories together with the corresponding values for *t*. According to these findings, the mean behavior efficiency rating for Group 1 is reliably lower than the means for Groups 3, 4, and 5, and the mean for Group 5 is reliably higher than the means for Groups 1 and 2. Group 5 is not reliably higher than Group 3 mainly because there were only six cases in the highest category. As shown in Table 2, the absolute values of the categories are all in the same direction. Thus there is a definite and direct relationship between measures of behavior efficiency derived from

TABLE 3

Differences* for Pairs of Behavior Efficiency Means and the Corresponding Values of *t*

Behavior efficiency means	Differences	Value of *t*
M_1 vs. M_3	8.97	2.24
M_1 vs. M_4	13.08	2.69
M_1 vs M_5	19.52	2.87
M_2 vs. M_4	10.78	2.80
M_2 vs. M_5	17.22	2.76
M_3 vs. M_5	10.55	1.18

*Differences between adjacent categories are not included. None are statistically significant at the three per cent level.

tests given during Training School residence and ratings of community social adjustment.

Presented together in Table 4 and Figure 2 are five group patterns or

TABLE 4

GROUP PSYCHOMETRIC PATTERNS INVOLVING TESTS USED IN ROUTINE CLINICAL PRACTICE

Test scores	Adjustment groups				
	1	2	3	4	5
Terman Vocabulary *Q's*	62	66	68	66	65
Stanford-Binet *Q*'s	68	69	71	71	69
SAR Performance *Q*'s	71	77	85	87	93
Pintner-Patteison *Q*'s	83	87	92	92	94

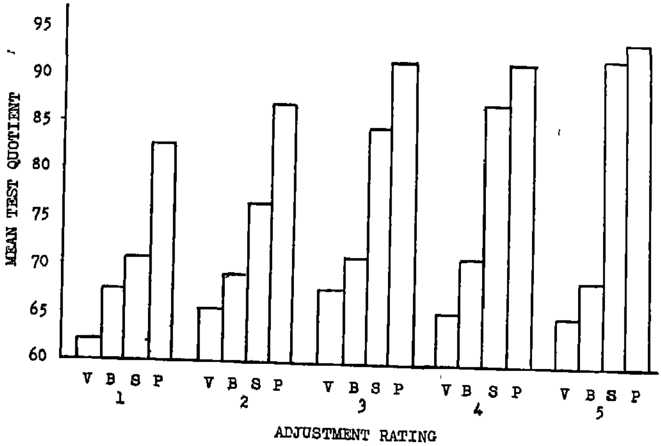


FIGURE 2

MEAN TERMAN VOCABULARY (*V*), STANFORD-BINET (*B*), *SAR* PERFORMANCE (*S*), AND PINTNER-PATTERSON PERFORMANCE (*P*) SCORES FOR EACH SOCIAL ADJUSTMENT CATEGORY

sets of group scores from the Terman Vocabulary, Stanford-Binet, *SAR* Performance, and Pintner-Patterson tests. The decimal places are eliminated in Table 4 to facilitate comparison. The following noteworthy facts are revealed:

1. The rank order of test values is the same in all five group patterns. Ranging from lowest to highest are the Terman Vocabulary, Stanford-Binet, *SAR,* and Pintner-Patterson.

2. The Stanford-Binet scores are remarkably constant from group to group, maximum variation being three points. Therefore functional intelli-

gence as measured by the Binet is not related to the behavior upon which level of social adjustment was judged.

3. Although Terman Vocabulary *Q*'s are consistently lower than Stanford-Binet *Q*'s, the differences are not statistically significant; they may be attributed to sampling errors.

4. Pintner-Patterson Performance means are consistently higher than *SAR* Performance means, and increase with adjustment ratings. Differences between Pintner-Patterson scores, even between Group 1 and Group 5, are not statistically significant at the five per cent level of confidence.

The comparison of *SAR* Performance with the Pintner-Patterson Scale as a whole reveals still another interesting finding. Going from Group 1 to Group 5, the discrepancies between the *SAR* and Pintner-Patterson scores steadily decrease—12, 10, 7, 5, 1. Evidently subjects with lower adjustment ratings derived their relatively high scores on subtests excluded from the *SAR* scale. In contrast, those with higher adjustment ratings made relatively higher Pintner-Patterson *PQ*'s from uniformly high subtest scores. Because of the heterogeneity of the Pintner-Patterson Performance items (insofar as power tests are mixed with speed, accuracy, and relevancy tests), final scores fail to be predictive of the behavior described by the evaluations of social behavior.

TABLE 5

MEAN EDUCATONAL SCORES FOR EACH SOCIAL ADJUSTMENT GROUP

Test scores	Adjustment groups 1	2	3	4	5
Stanford Educational Quotient	67	71	70	71	71
Stanford Reading Grade	4.9	4.7	5.0	5.0	5.3
Stanford Arithmetic Grade	4.4	4.5	4.7	5.2	5.6

Table 5 shows the relationship between academic achievement scores and ratings of social adjustment. Scores from the Stanford Achievement Battery in terms of educational quotients are given together with reading comprehension and arithmetic computation grade scores. The latter two scores are shown separately for they are measures of the most essential subjects in the academic program for mentally retarded children.

The Stanford Educational Quotients are practically identical; intergroup variation is within four points. Furthermore, each *EQ* varies not more than two points from the group's corresponding Binet *IQ* (Table 4). Although reading and arithmetic grade averages are practically the same throughout there is a decided trend toward a direct relationship between level of arithmetic attainment and social adjustment ratings. This is interesting in the light of the well-established clinical finding that mental patients find it difficult to do arithmetical computations.

DISCUSSION

In this group of 136 relatively homogeneous mentally retarded male subjects, measures of behavior efficiency extend over a wide range—from 9 to 29. Since this factor is found in varying degrees in a group of individuals who are not considered to be extreme in mental retardation or in severity of delinquency, then present findings are consistent with the contention that behavior efficiency is a general trait rather than a unique personality characteristic of such extreme clinical types as psychotics and prison recidivists. Subjects having the highest behavior efficiency scores on tests administered while at the Training School were among those making the highest social adjustment ratings six and one-half years after leaving the School or approximately nine years after the time of the tests. These subjects, as a group, were found to be supporting themselves and their dependents, holding good jobs, keeping clear of social and legal conflicts, etc. The reverse was true for those with the lowest behavior efficiency ratings: they were usually dependent on others, seldom or never worked, and were in constant social and/or legal entanglements. Decidedly, behavior efficiency is an ever-present factor which varies directly with level of social adjustment whether considered to be within or outside the normal range of behavior.

Pertinent to the present findings is Hamlin's study (8) on mentally retarded girls, an investigation "concerned chiefly with suggestions as to the significance of pattern scores." One-hundred-one high grade mentally defective girls were analyzed with respect to institutional adjustment and test patterns which were expressed as ratios between Stanford-Binet (1916) mental age and a performance age. Scores on the latter test were derived from a scale consisting of five items from the Pintner-Patterson Series. Four tests—Mare and Foal, Two-Figure Formboard, Casuist Formboard and Knox Cubes—were included because they were among the tests yielding lowest scores in subjects characterized by low performance quotients. The fifth test—Gwynn Triangle Formboard—was added arbitrarily. The criterion of social adjustment was based upon consideration for parole which was in turn contingent upon adjustment at the institution. It was found that none of the 21 girls with low ratios had ever been considered for or tried on parole, while 53 per cent of the 27 girls with high performance patterns had been given a parole trial. In addition, girls with particularly low performance patterns had, among other things, definite behavior disorders other than simple mental retardation.

Hamlin's method and procedure differed from those described here in several major respects. For example, test scores were related to predicted social adjustment[5] rather than to accounts of actual social behavior; the performance scales differed—only four items were the same in the two studies; the performance score was contrasted with the Stanford-Binet rather than with the Terman Vocabulary. The last difference is not serious in its consequence on final results since the findings in the present study would not have been modified if behavior efficiency had been defined by the disparity between *SAR* Performance and Stanford-Binet instead of the Terman Vocabulary test inasmuch as the discrepancies between Terman Vocabulary and Stanford-Binet were negligible. Despite differences between the two studies, the results are in substantial agreement.

What explanation can be offered to account for the relationship between behavior efficiency and social adjustment? In an attempt to answer the question, we advance the following theory: Each individual has a given inhibitory capacity which extends over a limited range. This ability—an inferred concept—has an important function in learning or relearning, particularly when these situations are highly complex. At a given level of intelligence, the ease of learning new complex act-sequences depends, among other things, on the ability to inhibit acts (responses) to irrelevant situations (stimuli). When the learning process is in reality a relearning situation (learning to make new responses to old stimuli), the capacity to inhibit plays an even more essential rôle. Here the new learning is held in abeyance until previously learned acts are completely inhibited. When inhibitory capacity is low, relearning is a slow and difficult process which takes place only when motivation is high.

Ratings on the *SAR* Performance test are quantitative expressions describing samples of behavior under rigid goal-striving or learning conditions. The behavior exhibited here is a function of many factors, including inhibitory capacity. When inhibitory capacity is inferior, the *SAR* score, based on time, errors, and moves, is low, since responding to extra test situations and coping with ensuing conflicts of action require added time and promote errors and extraneous activity. On the other hand, the Terman Vocabulary score discounts these factors, and the discrepancy between the two tests—behavior efficiency—offers some indication of inhibitory capacity. How individuals behave in the more complex personal and social goal-striving situations in everyday living depends in part on inhibitory capacity. Presumably, those who understand what is required of them and comply, without prolonged delay or difficulty, are those who are most likely to achieve a satisfactory social adjustment. Needless to say, mere ability to attain without numerous detours and conflicts, dictated social goals, does not insure adequate social adjustment. The goals to be striven for must be acceptable to society, i.e., not in conflict with the social norms.

---

[5] Hamlin's criterion of social adjustment assumes a high and positive relationship between institutional adjustment and subsequent social adjustment.

Formulations similar to the one presented here have been used descriptively by many clinical psychologists and psychiatrists. A problem behavior child may be described as "having practically no inhibitory powers" or an adult with a psychopathic personality as being "unable to inhibit his primitive instincts." Holsopple (10) has been more definite in his attempt to link capacity for inhibition and persistent social maladjustment. He suggested that inability to inhibit or unlearn social habits may be part of the cause of recidivism. Using mirror drawing tests on groups of reformatory inmates he presented some evidence in support of his contention. Insofar as the subjects studied here had been in conflict with community mores prior to Training School admission, poor social adjustment subsequent to parole may be considered recidivism. From this point of view, Holsopple's hypothesis and findings may be interpreted as being continuous with findings in the present study.

According to our findings, separate ratings on the Stanford-Binet, Pintner-Patterson, Terman Vocabulary, and Stanford Achievement tests are not related to the level of future social adjustment of mentally retarded male subjects. It does not follow, however, that *IQ*, *EQ*, reading age, and arithmetic age are unimportant factors in making for adequate social adjustment. Unquestionably, the range of adjustment and the placement of individuals in each of the five adjustment groups are a function of a multiplicity of causes including the general level of intelligence, the educational achievement, and others not considered here. What the range would be and which individuals would be in higher or lower adjustment groups if the general *IQ* and *EQ* were much higher (e.g., in the 90's) or much lower (e.g., in the 50's) is an unanswered question warranting further investigation.

Although separate test scores are not predictive of future social adjustment, when viewed collectively as patterns they have considerable significance for psychological analysis (2). Variations in test patterns (consisting of the Terman-Merrill Vocabulary, Revised Stanford-Binet, and Grace Arthur Performance Scale) of children now at the Wayne County Training School are related to level of institution adjustment (4, 5). Furthermore, the pattern shown in all five categories here (low Vocabulary, even, and low Binet, and high performance) is characteristic not only of the young mentally retarded predelinquent, but also of the mentally defective (14) and the adolescent offender committed to an industrial school (7).

SUMMARY

This study is primarily concerned with the relationship between behavior efficiency and the social adjustment of mentally retarded young men. Behavior efficiency, defined as the disparity between a speed-accuracy-relevancy performance test and a verbal power test, has been shown to be a significant concept in studying extreme clinical types such as psychotics and prison recidivists. To be useful in general psychological analysis, it must be shown to be a trait present in varying degrees in all individuals and related to personal-social adjustment at all points along the range. In preparing the material for study it was possible to examine separately, and in patterns, some of the usual clinical tests in relation to future social adjustment.

The subjects consisted of 136 young men who had once attended the Wayne County Training School. These boys were personally contacted by social workers between four and eight years after parole, for the purpose of evaluating social adjustment. On the basis of these data, each youth was rated on a five-point adjustment scale and the entire group was divided into five adjustment sub-groups. From tests administered while the subjects were at the Training School, the following mean scores were calculated: *SAR* Performance quotient (seven speed-accuracy-relevancy items from the Pintner-Patterson Performance Series), Terman Vocabulary quotient, Stanford-Binet (1916) Intelligence quotient, Pintner-Patterson Performance quotient, Stanford Educational quotient, and Stanford reading and arithmetic grades. The discrepancy between the *SAR* Performance scores and Terman Vocabulary scores was tested for relationship to social adjustment by applying t tests to group differences.

Means of behavior efficiency were found to vary directly with social adjustment ratings. According to the tests of significance applied, the mean for Group 1 was reliably lower, at the five per cent level of confidence, than the means for Groups 3, 4, and 5, and the mean for Group 5 was reliably

higher than those for Groups 1 and 2. When considered separately, all the other tests were found to be unrelated to social adjustment; when considered in patterns, several important findings were revealed.

The results are discussed in the light of similar findings by Hamlin (8) and other investigations concerned with psychometric patterns. A theory of capacity for inhibition is advanced to account for the relationship between behavior efficiency and social adjustment.

REFERENCES

1. Bijou, S. W. Psychometric similarities between habitual criminals and psychotics. *Delaware State Med. J.*, 1939, **11**, 126-129.
2. ———. The psychometric pattern approach as an aid to clinical analysis.—a review. *Amer. J. Ment. Def.*, 1942, **46**, 354-362.
3. ———. The measurement of adjustment by psychometric pattern techniques—its need in the selective service program. *Amer. J. Orthopsychiat.*, 1942, **12**, 435-438.
4. ———. An experimental analysis of Arthur performance quotients. *J. Consult. Psychol.*, 1942, **6**, 247-252.
5. Bijou, S. W., & McCandless, B. R. An approach to a more comprehensive analysis of mentally retarded pre-delinquent boys. (In preparation.)
6. Brown, J. F. Psychodynamics of Abnormal Behavior. New York: McGraw-Hill, 1940.
7. Glanville, A. D. Psychometric patterns of industrial school boys. *Delaware State Med. J.*, 1937, **9**, 91-94.
8. Hamlin, R. Test patterns of high grade mentally defective girls. *Proc. Amer. Ass. Ment. Def.*, 1938, **43**, 161-165.
9. Haskell, R. H. An organization for the training of high grade mentally defectives. *Proc. & Addr. Amer. Assn. Ment. Def.*, 1932, **56**, 252-270.
10. Holsopple, J. Q. The social adjustment of delinquents who are unable to inhibit old automatic perceptual responses. *J. Soc. Psychol.*, 1932, **3**, 91-96.
11. Jastak, J. Variability of Psychometric Performances in Mental Diagnosis. New York City, 1934.
12. ———. Psychometric patterns of state hospital patients. *Delaware State Med. J.*, 1937, **9**, 87-91.
13. Kephart, N. C., & Ainsworth, N. H. A preliminary report of community adjustment of parolees of the Wayne County Training School. *Amer. Assn. Ment. Def.*, 1938, **43**, 161-166.
14. Kinder, E. F., & Hamlin, R. Consistency in test performance patterns of mentally subnormal subjects. *Prof. Amer. Assn. Ment. Def.*, 1937, **42**, 132-137.
15. Lindquist, E. F. Statistical Analysis in Educational Research. Boston: Houghton-Mifflin, 1940.
16. Pintner, R., & Patterson, D. C. A Scale of Performance Tests. New York: Appleton, 1925.
17. Warren, H. C. Dictionary of Psychology. Boston: Houghton-Mifflin, 1934.

Wayne County Training School
Northville, Michigan

The Journal of Genetic Psychology, 1944, 65, 147-160

AN APPROACH TO A MORE COMPREHENSIVE ANALYSIS OF MENTALLY RETARDED PRE-DELINQUENT BOYS*[1]

Wayne County Training School, Northville, Michigan

S. W BIJOU AND B. R. MCCANDLESS

It is the purpose of this paper to analyze one aspect of the personality of mentally retarded, pre-delinquent boys in the hope of throwing some additional light on the behavior make-up of boys usually included in this category. The personality is thought of as an integrated set of traits, some rather fixed, some less variable, and some fluctuating Such a conception is consistent with J. F. Brown's formulation of personality (8, p. 13)

Brown speaks of three sources of personality (*a*) Characteristics of humans chiefly determined by heredity, under which he includes physique and reaction systems, the latter being the biologically determined "reflexes." (*b*) Characteristics of humans in which heredity and environment are co-determinant. This includes intelligence and its subtraits.[2] (*c*) Characteristics determined mainly by the environment (goals and emotional strivings are included here)

The authors do not wish, however, to be drawn into the controversy concerning the relative weights of heredity and environment in the determination of traits. Throughout the paper we shall speak, then, of (*a*) relatively constant, (*b*) moderately constant, and (*c*) variable traits of personality. Those traits referred to as relatively constant may be thought of as corresponding to Brown's first group of traits; those referred to as moderately constant correspond to his second group, and traits spoken of as variable fit in with Brown's third grouping.

Extensive studies on the subject of mental retardation and delinquency have furnished excellent insights into the areas of physical characteristics of these groups of individuals. Burt (9), Glueck and Glueck (11), and Anderson and Leonard (1), among others, have reported on health and

---

*Received in the Editorial Office on December 11, 1942

[1] From the Wayne County Training School, Northville, Michigan, Robert H. Haskell, M D., Medical Superintendent, Research Department, Thorleif G Hegge, Ph D, Director of Research and Education

[2] The authors believe that this category is far more extensive than Brown believes it to be Our point of view is maintained in the discussion that follows

physical make-up of delinquents Likewise a considerable amount of excellent material has been gathered concerning the third source of personality (goals and emotional strivings). There has been a keen appreciation of the importance of this trait source. Healy and Bronner's work (12), investigating the mechanisms of delinquent behavior, and regarding this behavior in terms of (*a*) its goals, (*b*) the emotional and situational learning processes responsible for shaping the behavior directed toward these goals, in the way it has been shaped, and (*c*) methods and probably chances of success in modifying these behaviors, has been a valuable contribution It has helped to open the field for understanding and retraining the delinquent, in terms of the experimental work done, for instance, in such areas as frustration and aggression, fixation and regression, etc.

Studies of the rôle of those personality factors which are moderately constant have concerned themselves chiefly with ascertaining the intelligence of delinquents as an etiological factor. Some authors have passed beyond this stage and have introduced such general concepts as hyperactivity, suggestibility, constitutional inferiority, and emotional instability in their consideration of the delinquent or mentally retarded. The quantitative data pertaining to these traits have thrown some light on this second aspect of personality, however they have not always been well-defined or carefully integrated among themselves or with the total personality, but vaguely set forth as etiological entities.

It is as an attempt to define and integrate a set of concepts within the area of this second, moderately constant, source of personality that the present study has been undertaken Our aim is to set up certain definite personality concepts, show how they are interrelated, how they may eventually result in a unified conception of this particular aspect of personality, and how they may be utilized in the better understanding of the personality of mentally retarded, pre-delinquent boys The instruments for the formulation of our concepts are psychometric patterns. Thus our terms will be different from those usually included in this category of personality, yet they will overlap in meaning with such frequently used concepts as constitutional inferiority, emotional instability, and verbal retardation.

The psychometric tests used in this study were the 1937 revision of the Stanford-Binet, and the Grace Arthur scale. An additional quotient was taken from the Binet (Form *L*) the vocabulary quotient, secured from the vocabulary list These three quotients, used together, constitute the psychometric pattern with which we shall be concerned. A psychometric pattern is the distribution of scores from any series of standardized psychometric tests, the tests hypothetically designed to measure different behavior areas. It is expressed in terms of its quotients, e g., *VQ* (vocabulary quotient) 80; *IQ* (intelligence quotient) 95; *PQ* (performance quotient) 110.

Considerable work has already been done on the meaning of patterns (4). Some investigators have sought distinguishing patterns for precisely defined clinical groups (such as psychotics, neurotics, prison recidivists, educational disabilities, etc), and have related the concept defined by the pattern to the major characteristics of the groups. Some have proceeded in reverse fashion by starting with widely divergent patterns and exploring other differentiating characteristics of the groups so separated Our intention is to use this existing material in addition to that gathered for our own analysis in amplifying our discussion.

SUBJECTS AND PROCEDURE

The subjects for our analysis were boys resident in the Wayne County Training School, an institution for high-grade mentally retarded children, boys and girls A combination of three factors is found to be present in the majority of commitments: (*a*) Intelligence quotient of 75 or below at the time of admission; (*b*) extreme school retardation; and (*c*) socially unacceptable behavior of a pre-delinquent type Low socio-economic status is generally associated with most of these cases.

The aim of the institution is to transform a child whose incapacity has been threatening social failure, into a self-respecting individual who can, with a minimum of assistance, maintain himself successfully in the community. This is accomplished first, by advancing him academically; second, by retraining him behaviorally, and third, by training him in skills and vocational habits so that he can take his place in normal, competitive society

The psychometric patterns for practically the entire boys' population were used in this study. The number of cases was 454. The average length of institutional residence of all children discharged from the Training School has been very little over three years Children range in age, upon admission, from 8 to 16 years, the average age being about 12 years. The average Binet *IQ* for the subjects used for this analysis was 69, with a range from 50 to 89 Approximately 80 per cent of the boys were white, 24 per cent of them the children of foreign-born parents The remaining 20 per cent were negro boys, most of them the children of parents who had come from the South

As mentioned, the patterns under consideration were derived from three tests: 1937 revision of the Stanford-Binet, the vocabulary section from Form *L* of the 1937 revision, and the Grace Arthur Performance Scale (2). The Arthur quotient was determined from a maximum growth age of 16 years; the other two tests used the maximum growth age of 15, as suggested in the Terman-Merrill manual (18) This difference in age divisor actually was negligible, as few of the subjects were more than 13 years old at the time the initial Arthur was administered The Binet administered nearest in date to this initial Arthur was used for purposes of completing the patterns, and was the test which supplied the vocabulary ratings

RESULTS AND DISCUSSIONS

By averaging the three sets of test scores, the following general pattern is found for the male population: *VQ* 68, *IQ* 69, and *PQ* 78. Thus the typical boy at the Wayne County Training School has a performance-skewed pattern with Binet and vocabulary quotients practically equal.

Because of the uniqueness of the Training School, one cannot compare this group and their patterns directly with other groups. But the pattern found is similar to that reported by other investigators for institutionalized, mentally retarded and/or delinquent children Glanville (10), for instance, studied an institutionalized group of 30 boys with severe behavior problems and found a pattern similar to the one found for the present population. His *Q*'s are higher, as emphasis for institutionalization for his group was severely anti-social behavior rather than mental retardation The pattern obtained by Glanville was· *VQ* 80, *IQ* 80, and *PQ* 97

Kinder and Hamlin (16), studying a group of unselected mentally subnormal children, report that "Among Letchworth Village subjects, a pattern based on superiority in the Pintner-Patterson over the Binet of 10 or more *IQ* points appeared three times as frequently as did the pattern showing a discrepancy in the opposite direction."

Johnson and Fernow (15) found substantially the same results with a large group of mentally retarded subjects at the Dixon State Hospital They used the Cornell-Coxe Performance Scale rather than the Pintner-Patterson

Although our group pattern is performance-skewed, it is obvious that not all of the boys' patterns conform to the general type. Breaking down our total group for further analysis, we find, as would be anticipated, that the distribution of types of patterns differs markedly from that of a normal population. The Grace Arthur Performance Scale was standardized so that 50 per cent of the quotients derived fell within five points of the 1916 revision of the Binet. Twenty-five per cent were above and below respectively. We assume this to be true for the 1937 revision because of the high correlation between the two scales. Table 1 shows the marked divergence of our group from the original Arthur distribution.

TABLE 1

			Per cent *WCTS*	 G Arthur
Group I	(Arthur low below	—6)	13	25
Group II	(Arthur equal	±5)	27	50
Group III	(Arthur high above	+6)	60	25

It is evident from these figures that our group of boys, when compared to a normal group, are not only inferior in intelligence, but a large majority of them are characterized by a performance-skewed pattern. The fact that 40 per cent of the group have either an even or a low performance pattern is still an important consideration from the point of view of pattern analysis. In the Grace Arthur normal group, there is a one to one relationship between the proportions of children having high and low performance patterns (25 per cent in each instance). In the Kinder and Hamlin group of unselected mentally defective children, it will be recalled that the ratio was one to three in favor of high performance patterns. In the present study on high grade mentally retarded, pre-delinquent subjects, it is one to five; in Glanville's study on delinquent, non-defective boys, it was one to six[3]. Thus both the defective and the delinquent child tend to have performance-skewed patterns but the tendency increases from the simple defective to the mentally retarded pre-delinquent to the non-defective delinquent. These findings are interesting and will be developed further in the course of the paper.

As shown here and as described elsewhere (4), characteristic test patterns have analytical significance. To continue our analysis, it will be expedient to fractionate our total group into homogeneous pattern subgroups. The average *VQ*, *IQ*, and *PQ* for each of the groups is shown in Table 2.

These figures show that our total male population may be thought of as consisting of a group of high grade mentally retarded boys having one of three major types of patterns. What personality characteristics may be ascribed to the members of each group? What do these differentiating pat-

---

[3] Personal communication.

TABLE 2

		VQ	*IQ*	*PQ*
Group I	(13%)	73	72	65
Group II	(27%)	67	67	67
Group III	(60%)	68	67	84

terns mean in terms of learning and adjustment? These questions summarize the two-fold purpose of the discussion which follows.

On the basis of past pattern studies, we advance the following hypothesis concerning some of the personality traits of the boys in the various groups:

The boys in Group I (low performance group) have:
1. Low behavior efficiency
2. Mild or no verbal learning interference
3. Functional intelligence below latent intelligence.

The boys in Group III (high performance group) have:
1. Adequate or good behavior efficiency.
2. Marked verbal learning interference
3. Functional intelligence below latent intelligence

The boys in Group II (even pattern group) have either:

A
1. Adequate behavior efficiency
2. Mild or no verbal learning interference
3. Functional intelligence equal to latent intelligence

or *B*
1. Low behavior efficiency
2. Marked verbal learning interference.
3. Functional intelligence lower than latent intelligence

Explanation of Terms

1. *Functional intelligence*—The *IQ* on the Stanford-Binet Scale
2. *Behavior efficiency*—This construct is defined by the discrepancy between two selected tests measuring behavior under prescribed contrasting conditions In clinical practice, one test has usually been a standardized performance test such as the Army Performance Scale (14) or the Grace Arthur (7), the other a verbal power test[4] such as the vocabulary section of the 1916 or 1937 revision of the Stanford-Binet. When the disparity is in favor of the performance rating, behavior efficiency is said to be high, when it is in favor of the vocabulary test, behavior efficiency is low Neither type of test mentioned is entirely satisfactory for defining behavior efficiency, but their use has been forced inas-

---

[4] Power tests are tests which gauge the "degree of an individual's ability in a given field, measured by the difficulty of tasks which he can perform without regard to time" (19).

much as they are the best available for this purpose The concept is adequately defined by the relationship between a performance test which is scored solely on speed, accuracy and relevancy, and a vocabulary test which is evaluated on qualitative and quantitative considerations. Until a scale is constructed for the specific purpose of measuring behavior efficiency, the regular clinical tests named above must suffice (5)

In exploring the meaning of this concept (i e, with what other measures of behavior it correlates), evidence has been found that it is similar to such terms as stability, the mental organization-disorganization continuum, constitutional make-up, etc

3. *Verbal learning interference*—This inferred concept describes the factor or condition causing verbal and reading retardation in subjects with relatively adequate learning ability in nonverbal tests It is one of the major causal factors in "reading disability" It does not as yet have a concise independent definition. Eventually this may be stated in terms of laterality, capacity for retroactive inhibition, or some similar structure or process In the Binet, a mixed test, verbal learning interference may be reflected to a marked degree
4. *Latent intelligence*—The theoretical Stanford-Binet *IQ* when behavior efficiency and verbal learning interference have been eliminated As yet there is no test that gives an independent measure of latent intelligence.

ANALYSIS OF GROUPS I AND III

According to the definition of terms, one outstanding difference between Groups I and III (performance low, and performance high, groups) is that Group III has a higher behavior efficiency rating than has Group I Studies on various clinical groups have indicated that among other things level of behavior efficiency is related to the level of personal-social adjustment (6, 13, 14) If this also applies to the present group of subjects, then the boys in Group III, or those with patterns similar to that of Group III, should make better institutional and post-institutional adjustments than those in Group I. With regard to institutional adjustment, patterns have been determined for cottage groups typified by different social adjustments

At the Training School, two honor cottages with rights of self-determination by the boys and extra, earned privileges, have been established. Boys showing unusually good behavior and judged stability, are allowed to apply for membership in these cottages, and are admitted provided the members of the cottage approve them Homestead Cottage is the more mature group of the two—entrance is more difficult, and higher behavior standards are set up for the boys in residence

> The boys in Cottage 7 are younger chronologically and are not expected to assume as full responsibility for their behavior as are the Homestead boys, the requirments for entrance into Cottage 7 are less rigid than those for entrance to Homestead, and the number of boys in the cottage is larger
>
> Cottage 5 is designed for aggressive boys on the assumption that these boys, who in other cottages have defied adult authority or bullied other children will, if placed together, neutralize or counteract one another's aggressive tendencies
>
> Cottage 4 is a strictly cross-sectional cottage, boys are placed in it as a usual part of their residence in the Training School, with no selection other than being in their early to middle 'teens As they advance in age and maturity they may be placed in one of the three cottages described above.

In terms of good institutional adjustment, the four cottages may be ranked as follows first, Homestead, second, Cottage 7; third, Cottage 4, and fourth, Cottage 5.

Test patterns for the population of each of the above cottages are shown in Table 3.

TABLE 3

	Number	*VQ*	*IQ*	*PQ*
Homestead	23	70	75	91
Cottage 7	42	67	74	85
Cottage 4	44	66	69	78
Cottage 5	31	67	68	71

Table 4 shows the percentage by cottage of the three possible types of psychometric patterns

TABLE 4

	Per cent *PQ* high	Per cent *PQ*=*IQ*	Per cent *PQ* low
Homestead	71	25	4
Cottage 7	63	25	12
Cottage 4	58	29	13
Cottage 5	45	30	22

The *PQ* differences between Homestead and Cottage 4, Homestead and Cottage 5, Cottage 7 and Cottage 4, and Cottage 7 and Cottage 5 are all statistically significant at or below the two per cent level, according to Fisher's formula for the significance of differences of small samples (17)

It can readily be seen from the first of the two sets of figures above that the patterns vary directly with the adjustment ranking of the four cottages, hence in accordance with previous findings there is a direct relationship between behavior efficiency ratings and institutional adjustment Evidently more boys with adequate or high behavior efficiency (high performance patterns) gravitate toward the honor cottages, more boys with lower behavior efficiency tend to find their way to Cottage 5.

The above finding is reinforced by the results of another recently completed investigation on a selected group of boys from this population (7). Thirty-five boys with high Arthur *PQ*'s were matched on Binet mental age and chronological age with 35 boys with low *PQ*'s This group was studied intensively to determine in what other respects the two groups differed Among other things, it was found that 74 per cent of the high *PQ* or high behavior efficiency boys were judged by those responsible for their training, to be better adjusted than their paired, low behavior efficiency mates

Our second expectation, relating to the differential post-institutional adjustment of boys with high and low behavior efficiency has been realized in a separate study by Bijou (5). On the basis of social workers' summaries obtained from direct contacts, the social adjustment made during an average period of six years following parole from the Wayne County Training School was studied. All boys were rated on a carefully constructed five-point social adjustment scale by two raters. Group patterns were ascertained for the five adjustment groups obtained by using the Terman vocabulary, the Stanford-Binet (1916 revision), and a performance scale especially devised for the analysis The most markedly performance-skewed (performance high) patterns characterized the group which had attained the highest type of social adjustment. Boys with the lowest behavior efficiency, on the other hand, made the least adequate community adjustments and were involved in at least one major or several minor difficulties

To be able to state, then, that the boys in Group I (low performance boys) have low behavior efficiency, and that those in Group III (high performance boys) stand high in this trait is significant and meaningful in terms of the actual and potential personal-social adjustment of those in each group where behavior efficiency is low, adjustment is less adequate than when it is high We do not, needless to say, lose sight of the fact that behavior efficiency is only one of the conditions determining adequacy of adjustment.

A second major difference between the boys in Groups I and III according to our formulation, is that those in III have marked verbal learning interferences while those in Group I have little or none of this characteristic

This differentiation can be made from consideration of the patterns when they are interpreted in the light of evidence which has been gathered relevant to verbal learning interference. The difference exists despite the fact that we know verbal learning interference influences Binet scores and that neither the average Binet *IQ*'s nor *VQ*'s of Group I differs significantly from those of Group III. The evidence to substantiate this point of view is again furnished by the study (7) in which boys from this population were matched on mental age and chronological age for the Binet, but differed widely in Arthur *PQ* level. The low *PQ* group (comparable to Group I) was found to stand significantly higher in reading achievement than was the high *PQ* group (comparable to Group III) Furthermore, the boys with high *PQ*'s not only scored lower in reading than their matched mates but their reading attainment was lower than their own arithmetic scores The low *PQ* group was variable in this respect Thus it is analytically meaningful to state that the boys in Group I differ from those in Group III not only in terms of the different degrees of behavior efficiency possessed but also in the different strength of verbal learning interference

Finally, it follows from the foregoing that Groups I and III are similar in that both show functional intelligence lower than latent intelligence (latent intelligence is defined as the theoretical Binet *IQ* with low behavior efficiency and high verbal learning interference effects eliminated). Since Group I has low behavior efficiency and Group III has high verbal learning interference, then by definition the boys in both groups have a lower functional intelligence than latent intelligence since both factors are assumed to have an adverse effect on Binet *IQ*. Thus, although superficially the end result is the same, the factors responsible for the functional intelligence-latent intelligence discrepancy are of an entirely different nature for each group

In summary, evidence has been presented to show that the boys in Groups I and III, although they differ little in functional intelligence, show rather different behavior make-up. They possess quantitative differences in several personality traits. Those in Group I are low in behavior efficiency and have little or no verbal learning interference, while those in Group III are high in behavior efficiency and have marked learning interference These findings provide a better understanding of these boys since it has been shown that behavior efficiency is directly related to level of personal-social adjustment while verbal learning interference is inversely related to learning to read, as well as to other forms of verbal learning.

ANALYSIS OF GROUP II

In the light of constructs developed above and applied to Groups I and III, Group II, the boys with even patterns, logically has two subdivisions. According to our hypothesis, one group (II*A*) should be composed of boys who are evenly retarded in most spheres of behavior and have little or a negligible discrepancy between functional and latent intelligence. On the other hand, those in the other group (II*B*) should have both low behavior efficiency and high verbal learning interference.

We pointed out earlier that for Group I (low *PQ* boys), among other things, the low performance pattern reflects low behavior efficiency; in Group III, the high performance pattern indicates high verbal learning interference. Accordingly, functional intelligence is lower than latent intelligence for both groups. Where low behavior efficiency and high verbal learning interference are present simultaneously, their combined action can very well produce the even patterns. In such instances the disparity between functional intelligence and latent intelligence would have to be marked. Thus it is possible that boys with even patterns, such as those in Group II, may have rather different behavior structure.

What empirical data can be brought forward to demonstrate that Group II is only superficially homogeneous, and in terms of the concepts described consists of two subgroups?

On the basis of suggestions from a previous study concerned with a similar problem (3), an examination of the subtests of the Stanford-Binet was made on the assumption that the operation of low behavior efficiency and high verbal learning interference might be revealed.

The Binet, being a collection of heterogeneous subtests, from the present point of view, permits the child with lowered functional intelligence to indicate something of his latent intelligence by his ability to pass certain "upper" items far removed from his Binet mental age. It would follow, and this is borne out clinically, that these items are usually relatively free from the demands required for measures of behavior efficiency and verbal learning interference. On the average, then, the child with a lowered functional intelligence is expected to show a wider difference between his mental age and "upper" ability than the child whose functional intelligence and latent intelligence are practically the same. A consequence of this line of reasoning is that the distribution of differences (upper age - mental age) should tend to be bimodal those evenly retarded (II*A*) will show smaller discrepancies than those with functional intelligence lower than latent intelligence (II*B*)

In order to evaluate the upper limits of Stanford-Binet tests, a separate score which for convenience we shall call the ceiling quotient, was calculated for each of the 80 subjects with an even Stanford-Binet–Grace Arthur pattern. This score was derived from the test value of the highest tests passed plus the age level of the group of tests immediately below the age level in which the highest test was passed (e g, the subject passed Test 2, at the 14 year level, with an age value of two months. The "ceiling basal" was regarded as year 13, so the mental age used for this ceiling quotient was 13-2) Ceiling quotient rather than ceiling age was used to cancel differences in age. The actual *IQ* from the test was subtracted from the ceiling quotient (*CQ*) This difference which was positive and ranged from 0 to +44 was used as a rough indication of the action of low behavior efficiency and high verbal learning interference on functional intelligence.

Casual observation of the frequency distribution of *CQ* - *IQ* differences reveals a definite tendency toward bimodality (i.e, there is one cluster of values between 8 and 13 and a second between 17 and 23) The general mean is 16 89, the mean for the lower division is 9.64, and for the upper division 22.83.

We have, then, from intra-test analyses of the Stanford-Binet tests, gathered some evidence to show that there are sub-groups within the apparently homogeneous even-patterned group, the first subgroup is one with relatively little upward scatter on the Binet, which indicates according to our hypothesis that relatively little verbal learning interference or behavior efficiency inadequacy is operating (II*A*) The second is a group where there is much upward scatter (II*B*), and verbal learning interference and behavior inefficiency are marked.

SUMMARY

The present analysis was made in an effort to further our understanding of the personality structure of mentally retarded pre-delinquent boys. More specifically it was concerned with differentiating and integrating those moderately constant personality characteristics which in J. F Brown's treatment of personality are spoken of as co-determined by heredity and environment. As a first approximation at this differentiation and integration four independent yet closely interrelated concepts were introduced · functional intelligence, latent intelligence, behavior efficiency, and verbal learning interference. Each construct was defined or was potentially definable by psychometric relationships (psychometric patterns). By investigating the patterns (consisting of the Terman-Merrill vocabulary, the revised Stanford-Binet, and the Grace Arthur Performance Scale) of all boys in residence at the Wayne County Training School, we found that mentally retarded pre-delinquents were a heterogeneous group. The largest proportion of the group had lowered functional intelligence, high behavior efficiency, and high verbal learning interference. About 13 per cent had lowered functional intelligence accompanied by low behavior efficiency and low verbal learning interference A similar proportion also had lowered functional intelligence, simultaneously with low behavior efficiency and high verbal learning interference The remaining boys had functional intelligence practically equal to latent intelligence without marked loss in behavior efficiency or undue verbal learning interference.

Not only were these factors defined and shown to be present in varying degrees for the subgroups of boys under consideration, but evidence was marshalled to show that they have analytical meaning and that they are important determining conditions in certain learning and adjustment situations For example, findings were produced which indicated that boys with high behavior efficiency tended to make better institutional and post-institutional adjustments than boys with comparable functional intelligence but low behavior efficiency; boys with high verbal learning interference were found to be inferior in reading achievement to boys with low verbal learning interference despite practical equality of functional intelligence.

Brown conceives the moderately constant personality traits as consisting only of intelligence and its subtraits. The position taken here is that from this source of the personality arises a set of interrelated traits like and including those described above.

The concepts which have been discussed and their interrelationships have been found to be exceedingly valuable in all phases of clinical practice In making psychological or educational analyses, in attempting to evaluate future institutional behavior, and in recommending or executing guidance programs, these factors should be taken into account. Needless to say, they must always be considered in conjunction with all the other known traits and conditions applying to the individual or group. Little is known, however, concerning the way in which the factors treated here influence, interact, and interrelate with those from the other personality sources. As an attack upon a segment of the total problem we are now attempting to analyze and manipulate the goals and emotional strivings (the personal-social habit structure) in a small group of boys with different constellations of the traits described here.

REFERENCES

1. ANDERSON, V. V., & LEONARD, C. M. A study of the physical condition of one thousand delinquents seen in court. *J. Crim. Law & Criminol.*, 1919-20, **10**, 82-89.
2. ARTHUR, G. A Point Scale of Performance Tests. New York: Commonwealth Fund, 1930.
3. BIJOU, S. W. The problem of pseudo-feeblemindedness. *J. Educ. Psychol.*, 1939, **30**, 519-526.
4. ———. The psychometric pattern approach as an aid to clinical analysis—a review. *Amer. J. Ment. Def.*, 1942, **46**, 354-362.
5. ———. Behavior efficiency as a determining factor in the social adjustment of mentally retarded young men. *J. Genet. Psychol.* (In press.)
6. ———. The measurement of adjustment by psychometric pattern techniques. *Amer. J. Orthopsychiat.*, 1942, **12**, 435-438.
7. ———. An experimental analysis of Arthur performance quotients. *J. Consult. Psychol.*, 1942, **6**, 247-252.
8. BROWN, J. F. Psychodynamics of Abnormal Behavior. New York: McGraw-Hill, 1940.
9. BURT, C. The young delinquent. New York: Appleton-Century, 1925.
10. GLANVILLE, A. D. Psychometric patterns of industrial school boys. *Delaware State Med. J.*, 1937, **9**, 91-94.
11. GLUECK, S., & GLUECK, E. T. One Thousand Juvenile Delinquents. Their Treatment by Court and Clinic. Cambridge, Mass.: Harvard Univ. Press, 1934.
12. HEALY, W., & BRONNER, A. F. New Light on Delinquency and its Treatment. New Haven, Conn.: Yale Univ. Press, 1936.
13. JASTAK, J. Variability of Psychometric Performances in Mental Diagnosis. New York City, 1934.
14. ———. Psychometric patterns of state hospital patients. *Delaware State Med. J.*, 1937, **9**, 87-91.
15. JOHNSON, A. P., & FERNOW, D. L. Comparison of results of Stanford-Binet and performance tests given at the Dixon State Hospital. *Proc. Amer. Assn. Ment. Def.*, 1939, **44**, 103-109.
16. KINDER, E. F., & HAMLIN, R. Consistency in test performance patterns of mentally subnormal subjects. *Proc. Amer. Assn. Ment. Def.*, 1937, **42**, 132-137.
17. LINDQUIST, E. F. Statistical Analysis in Educational Research. Boston: Houghton-Mifflin, 1940.
18. TERMAN, L. M., & MERRILL, M. A. Measuring Intelligence. Boston: Houghton Mifflin, 1937.
19. WARREN, H. C. Dictionary of Psychology. Boston: Houghton-Mifflin, 1934.

Wayne County Training School
Northville, Michigan

The Journal of Genetic Psychology, 1944, 65, 161-175.

AN EXPERIMENTAL AND CRITICAL STUDY OF THE INTELLIGENCE OF NEGRO AND WHITE KINDERGARTEN CHILDREN*

Minneapolis Public Schools

Fred Brown

A Introduction

The urgent need for additional manpower in war industries has probably done more to focus attention upon the potentialities of the Negro than has been accomplished by legions of psychological and sociological research workers. Whatever else it may ultimately accomplish, total war has compelled us to utilize the human components of our society in the most immediately practical manner We are slowly thrusting aside academic considerations, prejudices, and socio-economic barriers in our efforts to utilize human capacities and resources wherever they are needed Deeply-rooted and unconsciously buttressed beliefs in the innate superiority and inferiority of peoples are just beginning to yield to the stubborn attrition of brute facts The grudging admission of small groups of Negroes to some flight-training schools, machine shops, and other jealously white-monopolized fields may, in time, bring about a revision of many of our fallacious conceptions of Negro intelligence and stability From the point-of-view of scientific research, studies of Negro intelligence leave one with the feeling that investigators have been measuring the *effects* of a certain culture upon the mental-test performance of individuals who are characterized by a heavily pigmented skin rather than the possible correlation between skin pigmentation and intelligence. Conditions affecting test scores have not been rigorously controlled in most of these studies, especially so basic a factor as socio-economic status. This variable is an especially difficult one to control because of the restrictions placed upon Negro employment in all but the lowest occupational levels It is very likely therefore that the manner in which Negroes are distributed among the various occupational levels may reflect lack of opportunity for diversity of occupational choice rather than limitations in general intelligence and specific aptitudes. If this assumption is correct we may not be entirely justified in assuming that White occupational classification scales

---

*Received in the Editorial Office on December 28, 1942.

permit the same conclusions with regard to the Negro's socio-economic status and intelligence as they do for the White population. Furthermore, even though we assume that dining car waiters, for example, represent persons of relatively good social and economic position among Negroes and that the White equivalent in terms of prestige and income would be roughly equivalent to the semi-skilled occupational class (Group V), we ought not to be surprised if the children of these waiters fail to score as well on tests of intelligence as do White children of comparable socio-economic status. A theoretical resemblance between two social groups in terms of socio-economic status as indicated by ratings based upon occupational rating scales does not necessarily demonstrate a real equivalence in cultural drive, motivation, and aspirational level[1] If we discovered that Negro children of semi-skilled fathers obtained mean *IQ*'s which were similar to those obtained by White children of similar socio-economic level this might point to the truth-value of the Negro occupational designation. Such an occurrence might also be interpreted as a confutation of the presumed intellectual inferiority of the Negro school child. If, on the other hand, such a group of Negro children scored either above or below the mean for Whites we must assume that the nominal socio-economic rating assigned this group was probably erroneous and/or that other factors must be taken into consideration in dealing with this problem. The problem of socio-economic status as it affects the Negro remains stubbornly insoluble so far as an accurate equating of groups is concerned as long as occupational opportunities are not the same for both Whites and Negroes In this study we have compared the test scores of Negro and White kindergarten children in an effort to determine how they compare at a young age when both groups are relatively unaffected by factors resulting from racial and cultural differences.

B THE NEGRO COMMUNITY

Negroes in Minneapolis are a small group comprising less than one per cent of the total population. Their presence has never led to class riots Their residences are more or less involuntarily segregated in certain sections of the city by the pressure of public opinion. Their living conditions are characteristic of those found in other large cities Here illiteracy among Negroes is relatively small, since 1.7 per cent of the Minneapolis Negro population over 10 years of age is illiterate and 16 3 per cent of the Negro population of the United States over 10 years of age is illiterate Even this low rate of illiteracy among Minneapolis Negroes would be lower were it not for the older people, since 44.5 per cent of Negro illiterates are over 65 years of age. Also, 55.3 per cent of the males are found in domestic and personal service and 83 0 per cent of females belong in this category. "The remaining male Negroes are porters, janitors, waiters, barbers, laborers, and unskilled workers in manufacturing industries and in street and railroad construction A very small percentage of the Negroes are doing clerical and professional work" (22) Also, 74.5 per cent were born in other states, with Missouri, Kentucky, Kansas, Tennessee, Iowa, and Texas contributing their contingents in that order.

---

[1] This may be particularly true for the Negro professional and semi-professional groups where assumed equivalence with Whites is apt to be fallacious For the Negro in general it may be said that improvement in socio-economic status is often saltatory and does not reflect a family history of intellectual traditions This would seem to be especially true for the professional class

C. SUBJECTS

Three hundred forty-one native White, second generation kindergarten children of Scandinavian extraction were compared with 91 Negro children attending the Kindergartens of the Minneapolis Public Schools The mean chronological age of the White subjects was 69.51 months, σ 3 93 The mean Negro age was 69 15 months, σ 3.47 The critical ratio was 0 48, indicating that both groups were equated for chronological age The White group consisted of 137 males and 204 females There were 50 Negro males and 41 females. The selection and distribution of our White subjects was based upon the percentage of gainfully employed adults in Minneapolis classified according to the Goodenough-Anderson Scale of Occupations (1930 census data) (10, 13) A comparison of occupational levels within each group by means of the *t*-test showed a general tendency for extreme levels to be more highly differentiated than adjacent levels. A marked relationship between level of occupation and level of intelligence was also noted As a further check an analysis of variance was made to determine whether the variation among occupational levels was greater for White males alone and White males and females combined than the variation within individual occupational levels. This was found to be the case without exception A test for homogeneity of variance (to determine whether the variances within levels were true estimates of the total variance for each group) showed homogeneity of variances within levels. Hence the significance of the variance ratios were found to be attributable to the differences among the means of individual levels (4). This rigorous treatment of the White data provided us with a reliable yardstick for the determination of the approximate socio-economic status of our colored subjects. A study of paternal occupations among the colored group showed that 63.7 per cent were listed as janitors, porters, waiters, laborers, and *WPA* workers, 14.5 per cent were listed as tailor, statistician, postal clerk, grocer, mail clerk, mechanic, engineer (janitor?), draftsman, stenographer, and clergyman, while 21.8 per cent were unlisted. But we may assume, upon the basis of Minneapolis Negro occupational samplings and the fact that school records generally contain certain data on the non-transient professional and semi-professional classes, that many fathers in this group probably fall into the semi- and un-skilled categories. Homogeneity of variance among the Negro (1.299, $p<5$) indicated that they came from the same population. Variance between Negro and White subjects was also found to be homogeneous.

D. Procedure

All children were given individual Stanford-Binet tests, Form *L*, by their kindergarten teachers. These teachers were trained to administer and score the tests as part of the Minneapolis Kindergarten Binet Testing Program. Hilden (14) obtained a correlation of .89 with a probable error of estimate of 4.7 *IQ* points upon retesting 274 teacher-tested children after a lapse of one year. Brown (6) re-tested 201 such children after a mean interval of 2.62 years and found a correlation of .86 with a probable error of estimate of 5.6 points. These results agree with retest findings of other investigators and demonstrate that highly trained and certificated teacher-testers at the kindergarten level are able to administer a reliable Binet test to their pupils. Details of the project have been published elsewhere (3, 4, 5, 6).

E. Results

1. *IQ*

The mean *IQ* of the White children ($N=341$) was 107.06, σ 13.1. The mean *IQ* for the total Negro group ($N=91$) was 100.78, σ 12.9. The critical ratio (*Diff.*/σ *diff.* was 4.13, indicating with certainty that the difference between means was not due to chance. It is interesting to note that there is virtually no difference in the sigmas of the two groups and that these sigmas agree with those reported by Terman and Merrill for this age group (25).

2. *Sex Differences*

Table 1 shows that Negro males make somewhat higher scores than do Negro females but the difference is not statistically significant. White females are superior to Negro males whereas White males tend to score higher than Negro males, the significance of the difference being somewhat doubtful. Both White males and females are definitely superior to Negro females. Whether this results from the factor of rapport between Negro girls and White female examiners remains to be determined. There is no sex difference in intelligence between White males and White females.

TABLE 1

SEX DIFFERENCES IN STANFORD-BINET IQ'S OF NEGRO AND WHITE KINDERGARTEN CHILDREN

	N	*M*	*σ*	*M1 - M2*	*D/σ* diff
Negro males	*50*	*103.02*	*11.91*		
a vs Negro females	41	98.0	13.55	5.0	1.85
b vs White males	137	106.8	12.52	3.8	2.27
c. vs. White females	204	107.2	12.85	4.2	2.80
Negro females	*41*	*98.0*	*13.55*		
a vs. White females	204	107.2	12.85	9.2	6.97
b vs White males	137	106.8	12.52	8.8	4.88
White males	*137*	*106.8*	*12.52*		
a. vs White females	204	107.2	12.85	0.4	.01

3. *Socio-Economic Status*

The correct interpretation of Negro socio-economic status is, as we have stated before, a difficult task because of the inequalities in occupational opportunities. It would appear to be methodologically feasible however to compare a social group of vaguely defined socio-economic status (such as the Negro) with another group sharing the same culture and separated into experimentally determined occupational levels (such as our White group.) While we most emphatically point out that this technique does not establish the absolute socio-economic identity of the colored group in terms of the social and cultural implications of such status in our culture, it should nevertheless enable us to surmise the *probable* position of our Negro group in terms of at least one variable, namely intelligence. Inasmuch as the relationship between socio-economic status and intelligence is a positive one for employed white adults we might assume that the same relationship holds for the colored group and that comparisons of mean *IQ*'s of one group with the other may give us a knowledge of the true socio-economic level of the colored group. This approach, while theoretically defensible, can be entirely valid only in those instances where the vaguely defined group is not hampered in its social and economic adjustments. In the case of the Negro, because of restrictions operating to limit social and economic mobility, any such generalization can be made only with reservations, even though most of the Negro children used in this study come from parents who are engaged in what would be regarded among those for whom color is no bar, as semi-skilled and unskilled occupations. We do not know whether Negroes who are so classified would be capable of attaining a higher position on the socio-economic scale in a free society or whether the positions they actually occupy represent endowment-limited ceilings.

The data in Table 2 show that the variances are homogeneous, demonstrating that our subjects come from the same population. The difference between the variances of Negro males and Group VII White males is significant, but since the difference between the means is not statistically significant a more rigorous treatment of the data in this region was not indicated. We have only computed variances in those cases where the number of subjects was small. It should be pointed out that the critical ratio of 2.26 between Negro females and Group VII White females is not statistically significant since the t-test revealed a p of $> .05$.

The comparison of mean Negro *IQ*'s with White groups broken down into separate occupational levels brings out some interesting and provocative findings. Table 2 shows that Negro males are inferior to White males of professional and semi-professional paternity at the kindergarten level. It is highly significant that *our Negro subjects are not inferior in tested intelligence to Whites in Group III (Clerical, retail business, and skilled trades!)* The slight but statistically insignificant difference at this level vanishes almost entirely at levels below this point. In view of the fact that a majority of our Negro subjects have fathers who may be classed as semi- and unskilled workers this is all the more significant, as it may suggest the fallacy of using White occupational norms for Negro workers. Our studies with White subjects (4) reveal no such results.

Negro females make significantly lower scores than White females at every level except VI and VII. The colored subjects as a whole are similar in socio-economic status to White semi- and unskilled labor classes. As this is what we might be led to anticipate solely upon the basis of Negro occupational listings (noting however that the low female scores are responsible for the total group results) we are justified in making the follow-

TABLE 2

Comparison of Negro Males and Females with White Males and Females of Known Socio-Economic Status (Minnesota Occupational Scale)

	N	*M*	σ	*M*1 - *M*2	*F* ratio (Variances)	*CR*
Negro males	*50*	*103.0*	*11.91*			
a) vs White males I	14	114.6	12.45	11.4	1.154 p>5	3.62
b) vs White males II	16	115.0	8.94	13.8	1.698 p>5	5.36
c) vs White males III	25	107.7	13.24	4.5		1.53
d) vs White males V	60	104.5	12.20	1.3		0.64
e) vs White males VI	7	101.8	10.88	1.4	1.048 p>5	0.32
f) vs White males VII	15	101.2	8.12	2.0	4.897 p<1	0.81
Negro females	*41*	*98.0*	*13.55*			
a) vs. White females I	11	109.0	11.33	11.0	1.331 p>5	2.97
b) vs White females II	17	113.7	16.97	15.7	1.628 p>5	3.65
c) vs. White females III	57	110.2	11.32	12.2		5.82
d) vs. White females V	86	106.2	11.25	8.2		4.35
e) vs White females VI	18	96.5	15.64	1.5		0.37
f) vs. White females VII	15	105.4	11.35	7.4	1.360 p>5	2.26
Total negro	*91*	*100.0*	*12.91*			
a) vs Total White I	25	112.1	12.30	11.3		4.03
b) vs. Total White II	33	114.4	13.03	13.6		5.15
c) vs Total White III	82	109.4	12.42	8.6		4.52
d) vs Total White V	146	105.5	11.68	4.6		2.78
e) vs Total White VI	25	98.0	14.66	2.8	1.329 p>5	0.86
f) vs Total White VII	23	105.7	10.58	4.9		1.88

ing assertion: *Negro children at the kindergarten level in the Minneapolis Public Schools are not inferior to White children of nominally similar socio-economic status in general intellectual capacity as measured by the Stanford-Binet Scale, Form L. Negro males compare favorably with White males of higher socio-economic status than their own.*

PREVIOUS STUDIES

Our findings differ from those reported in the literature on the intelligence of young Negro children. Koch and Simmons (16) tested seven year old Negro children on the Myers Pantomime and Detroit First Grade tests and found that 78 per cent of the Whites surpassed the Negro median. Graham (12) compared the scores of Negro children in Atlanta, Georgia, with national norms on the Otis Group Intelligence Scale. Differences between Negroes and Whites ranged from 3.10 to 9.09 points from ages 8 to 11 inclusive, favoring the Whites. The difference between 7-year-olds was only 35 points. Differences favoring Whites on the Advanced Examination (10 to 14 years) ranged from 11.01 to 25.00 points. Sunne (24) compared 110 White with 32 Negro children, using the Terman and Yerkes-Bridges Revisions of the Binet tests. The mean *IQ* of the White group at age 6 was 100. The mean Negro *IQ* was 95. Strachan (23) compared 7,474 White and 298 Negro children. The median *IQ* of the White kindergarten children was 105.7 and of the Negro 96.79, and 44.61 per cent of his Negro children fell into the 56-95 *IQ* range as compared with only 20.12 per cent of the Whites. Klineberg (15) administered the Stanford-Binet Scale (1916 revision) to Negro children whose parents had migrated from the south and who had lived in the north from one to over four years. The average *IQ* of migrants of less than one year of residence was 81.4; one to two years, 84.2; two to three years, 84.5; three to four years, 85.5; over four years, 87.4. New York born Negro children earned an average *IQ* of 87.3.

Sunne (24) found that 7-year-old Whites obtained a mean *IQ* of 96 as contrasted with 88 for Negroes. At age 8 the White mean was 93, Negro 88; age 9, 93 for Whites, 85 for Negroes, at age 10 the average White *IQ* had dropped to 90, the Negro *IQ* to 79. Garth, Lovelady, and Smith (7) tested 2,006 Negro children in Dallas and Oklahoma in Grades IV to IX on the Otis Classification and obtained an average *IQ* of 77.9. Peterson and Harrelson (20) used the Pressey, Otis, Haggerty, and Myers tests with groups of from 32 to 734 Negro children. They claim to have obtained a relatively higher type Negro than White, but the findings revealed an average Negro *IQ* of 75 and 80.83 per cent of Whites were more efficient than the median of Negro ability. Only 15 to 18 per cent of the Negroes reached the median for Whites. Garth and Whatley (9) found a median *IQ* of 75 for 1200 Negro children on the National Intelligence Test. McGraw (19) administered the Buhler "Babytests" to Southern Negro and White infants with the result that White infants proved superior to Negroes (as they also did according to height and weight norms). Bruce (7) examined 521 White and 432 Negro children in the rural South. The average Kuhlmann-Anderson group test *IQ* for White children (mean *CA* 9-6 years) was 88.10. Mean for Negroes (mean *CA* 9-9 years) 71.80. On the 1916 revision of the Stanford-Binet scale White children ($N=87$) obtained an average *IQ* of 90.07, σ 15.66. The Negro subjects ($N=72$) obtained an average *IQ* of 76.33, σ 14.22. The Grace Arthur Performance battery, using the same *N*'s, resulted in an average *IQ* of 94.21, σ 22.11 for Whites, 77.10, σ 14.08 for Negroes. Critical Ratios for these differences were all above 5.00. Boynton (1) concludes ". . . most studies point so definitely to the superiority of the average White over the average Negro that it would seem justifiable to assume this condition to exist, at least until such time as definite *objective* confutations can be brought forth by objectors to this Hypothesis."

DISCUSSION

It does not suffice to show that the average *IQ* of Negro children is lower than that of Whites. If we reject the myth of superior and inferior races we must probe deeper into the causes for these rather consistently observed differences. It is imperative that we ascertain whether such differences are attributable to nature or nurture. In doing so we have no intention of discrediting previous investigations, weak as many of these have been methodologically, but in these times the psychological research worker who is interested in social issues must frequently advance beyond his refined statistical techniques and inquire into the qualitative factors which may be responsible for his results. It is these very factors which are often disregarded by those who make use of scientific data in order to bolster reactionary prejudices.

Our findings have already shown that Negro children, residing in a community wherein the literacy level is higher than that found for Negroes in the United States, make higher score on the Stanford-Binet Scale than is customary for children of this group elsewhere. We have shown that these children are as intelligent as Whites of approximately the same occupational level and we have pointed out the well-known fact that the Negro is grossly hampered in his choice of occupations. Sunne's findings (24) show a drop in the Negro *IQ* from 95 at age 6 to 79 at age 10 Bruce (7) asserts "From 6 through 12 years of age, the Negroes show no greater decline in *IQ* than the Whites." When we consider the initially low *IQ*'s of her subjects and the stultifying environment in which they live this is not to be wondered at Our own experience with Negro children in the Minneapolis Public Schools indicates that such children make progressively poorer test scores as they grow older until the average *IQ* for Negroes at the sixth and seventh grades approaches the scores mentioned by other investigators School principals in Negro districts have told the writer that Negro children seem to grow duller as they grow older. It might be assumed therefore that mental development in the case of the Negro reaches its maximum at a relatively early age Such a hypothesis would be more tenable if he shared the social and economic advantages of the White. This he obviously does not do We are forced to search for another explanation which will rest upon the objectively verifiable facts of the Negro's social and economic milieu and which will also possess psychological validity.

Gordon (11) tested English Canal-boat children and found that 76 older children earned an average *IQ* of 69 6 whereas younger children tested at the average for English children. He found a correlation of —.755 between intelligence and chronological age. This finding, interestingly enough, did not lead him to the conclusion that Canal-boat children are innately inferior or reached maximal mental development much earlier than did land-based children Instead, these results were advanced in support of the theory that intelligence deteriorates in the absence of external stimulation.

Given equal opportunity to attend public schools in the North and therefore exposed to educational stimulation (unlike Gordon's subjects), why has the Negro school child failed to reflect this advantage and why does his *IQ* seemingly deteriorate? If we point out that he almost always resides in the poorest sections of our cities and that his schooling in many cases is not comparable in quality to that of White children we would expect him to score at approximately the same level as do Whites in his neigborhood The average *IQ* for Minneapolis Kindergarten children is 107.7 (2). The average *IQ* for White children in a Negro neighborhood and in the same school wherein most of our subjects were tested is 104. A comparison of percentiles for Negro and White children in the same school reveals that the Negro *IQ* lies in the 27th percentile whereas the White average in the same school lies at the 35th percentile. Nevertheless these scores are more alike than would be the case if we compared these Whites with Whites elsewhere in the city.

The theory which we wish to advance in this paper involves emotional factors in the life-development of the Negro which are not encountered in other cultural groups in America. We have suggested in a previous paper (2) that the variabilities of the *IQ* at various age-levels from two years to post-adolescence follows a rather consistent pattern not due to chance From two to six years the standard deviations of the *IQ* gradually decline until the lowest variability (12 5) occurs at age 6 Beyond this point the variabilities gradually grow larger until a maximum is again attained at the adolescent period We have interpreted this phenomenon to mean that home training up to this age involves inhibitions and restraints imposed from above in an effort to divert aggression and hostility into socially conventional patterns of behavior. When the child enters school he is exposed to social and intellectual stimuli which encourages differentiation This phase of the process is reflected in increasing variabilities. Since the standard deviation of the Negro *IQ* in our group is almost identical with that of the total White group and Terman's standardization population (σ 12.91 and σ 12.30) we may assume that this process of conformity patterning is very similar for both groups. We must ask whether increasing differentiation after the age of six years characterizes Negro development, since we assume that such changes in variability are culturally determined If opportunities and incentives for differentiation are withheld from the Negro child, is it not likely that this process of differentiation fails to occur to the same degree as for Whites? This problem lends itself to experimental verification. What looks like deterioration of the Negro *IQ* may in reality be a hemming-in of intellectual development and/or efficiency which strongly suggests the operation of a constrictive factor in Negro intellectual development. It is our theory that such a process of constriction or "pinching off" takes place and accounts for much of the Negro's personality and intellectual development. This would lead to a studious regard for outward forms of behavior characterized by set conformity to White norms, an inhibition of spontaneity in verbal behavior, and sharp reduction in the level of aspiration Such patterns of behavior have their counterpart in frustration experiments carried out with animals

There are a number of cogent considerations which lend support to such a theory The Negro attempts to adapt himself to a dominant culture corporation in which he holds very little stock and has a proportionately small voice in determining policies (the Poll-tax situation illustrates this dramatically). His level of aspiration is adjusted to his opportunities and there is little motivation for superior performance which will only lead to severe frustration later in life. The Negro child realizes (consciously or unconsciously) that his vertical mobility is more restricted than is that of almost any other group in this country. In the Southern states he is expected to assume an attitude of submissiveness toward Whites. Signs of aspiration higher than that tacitly permitted him by Whites is construed as arrogance and presumption If he should strive for White status he is almost certain to encounter bitter disappointment which is likely to lead to neurotic symptom formation.[2] Lemkau, Teitz, and Cooper (17) made a Mental Hygiene study of 3337 cases in the Eastern Health District of Baltimore, supplementing their findings with National Health Survey material. They state:

> In the "nervous" and psychoneurotic groups, Negroes furnish but 15 and 14 per cent of the cases respectively The proportion in the small cases with neurotic traits corresponds almost exactly to the representation of the Negro race in the total population of 20 years of age and over, which is 21 5 per cent This distribution pattern, a disproportionally small percentage of Negroes, is seen to prevail in most of the diagnostic subgroups among the psychoneuroses

Upon the basis of an "index of case finding" they further conclude ". . Our lower rate for adult neurotics among Negroes corresponds to the actual facts" This probably arises out of his sharp goal-reduction and would appear to offer further support to the theory of a constriction mechanism which enables the Negro to attune himself to a culture which creates a frustration trap by ostensibly offering freedom of opportunity for all peoples regardless of race or creed, but at the same time prohibits the Negro from implementing these concepts operationally

Intelligence tests (whether individual or group tests) presuppose a certain degree of articulateness on the part of the examinee We believe that the Negroes inhibit verbalizations among Whites as one of the overt manifestations of the constriction process. Klineberg (15) has said that the

---

[2] The author has in his files the case of a Negro adolescent who developed paranoid tendencies as a result of severe status conflict

Negro's language equipment is poor He cites no experimental evidence in support of this statement, but if we accept it at its face value it would be necessary to point out that a sharp distinction must be made between language facility as measured by White test standards (which, upon the basis of intelligence test results would necessarily be low) and language as a vehicle for the expression of attitudes. Since language reveals attitudes, and forbidden attitudes may attract hostility and/or aggression from the dominant Whites, discretion outweighs valor in the development of language habits Limiting one's self to "yea" and "nay" reduces threats to personal security while it concomitantly leads to linguistic atrophy through disuse. The effects of this verbal inhibition, as the need for it increases with more life experience, upon measured intelligence needs no further elaboration That it is very likely a logical explanation of *IQ* deterioration must also be given careful consideration.

SUMMARY

Three hundred forty-one native White, second generation kindergarten children were compared for intelligence with 91 colored kindergarten children of the same chronological age. All children were given the Stanford-Binet Scale, Form *L* Socio-economic status of the White subjects was accurately determined and rigorously analyzed by the analysis of variance method Socio-economic status of the Negro subjects was carefully scrutinized and interpreted in terms of the community pattern for the Negro Methodological assumptions underlying a comparison of White and Negro subjects at each of six occupational levels were discussed The following results were presented·

1. The mean White *IQ* was 107.06 as compared with a mean Negro *IQ* of 100 78. The variances were particularly identical

2. Negro males make somewhat higher scores than females, but the difference is not statistically significant.

3. White females are superior to Negro males, but the evidence does not clearly demonstrate White male superiority over Negro males Both White males and females are superior to Negro females

4. Variances for White and Negro *IQ*'s are homogeneous at all occupational levels, indicating that the subjects come from the same population.

5 Comparison of intelligence at various occupational levels reveals that Negro males are not inferior in tested intelligence to White subjects in Group III (Clerical, retail business, and skilled trades)

6 Negro females score lower than White females at all occupational levels except VI and VII.

7. The total Negro group resembles the White semi- and unskilled labor class.

A theory of developmental constriction based upon clearly defined cultural factors in the life of the Negro is advanced as an explanation of the apparent deterioration of intelligence among Negro children with increase in age This theory seems to account for the consistently lower *IQ*'s of Negroes in comparison with Whites without assuming racial inferiority.

References

1 Boynton, P L Intelligence, Its Manifestations and Measurement New York Appleton, 1933.

2 Brown, F. Variability of the *IQ* in relation to age on the revised Stanford-Binet scale *J. Genet. Psychol* (In press)

3 ——— The Minneapolis kindergarten Binet testing program an experiment in applied psychology *J Consult Psychol.*, 1942, **6**, 142-148

4. ——— A comparative study of the intelligence of Jewish and Scandinavian kindergarten children. *J Genet. Psychol* (In press.)

5. ———. Mental hygiene in the kindergarten. *Ment Hyg* (In press)

6 Brown, F., & Challman, S. A. Binet testing by kindergarten teachers as a mental hygiene measure. *Sch & Soc*, 1940, **52**, 668-672.

7. Bruce, M. Factors affecting intelligence test performance of Whites and Negroes in the rural South. *Arch. of Psychol*, 1940, No 252, Pp 99

8. Garth, T. R, Lovelady, B. E, & Smith, H. W. The intelligence and achievement of southern Negro children. *Sch. & Soc*, 1930, **32**, 41

9. Garth, T R, & Whatley, C. A The intelligence of southern Negro children *Sch & Soc*, 1925, **22**, 501

10. Goodenough, F. L., & Anderson, J. E Experimental Child Psychology. New York Century, 1931

11. Gordon, H Mental and scholastic tests among retarded children London: Bd. of Educ, Educ Pamphlet No 44, 1923.

12. Graham, V T. Health studies of Negro children I Intelligence studies of Negro children in Atlanta, Georgia Printing Office, Washington, D C, Reprint No. 1127, Pub Health Reports, 1926, **41** (P 2759)

13 Green, J., Berman, I., Traxler, M R., & Paterson, D G A manual of selected occupational tests for use in public employment offices. *Univ of Minn. Press*, Bulletin Series, **2**, No 3, Employment Stabilization Research Institute, 1933

14. Hilden, A. H Training kindergarten teachers to test their own pupils on the Stanford-Binet scale. *Sch. & Soc*, 1938, **48**, 123-124

15. Klineberg, O. Negro Intelligence and Selective Immigration New York Columbia Univ Press, 1935. Pp 66

16. Koch, H. L, & Simmons, R. A study of the test performance of American, Mexican, and Negro children *Psychol. Monog*, 1926, **166** (P 35)

17. LEMKAU, P., TIETZ, C, & COOPER, M Complaint of nervousness and the psychoneurosis *Amer J. Orthopsychiat.*, 1942, **12**, 214-223

18. LONG, H H The intelligence of the colored elementary pupils in Washington, D C *J Neg. Educ*, 1924, **3**, 205-222

19 MCGRAW, M. B A comparative study of a group of Southern White and Negro infants *Genet Psychol Monog*, 1931, **10**, 1-105.

20 PETERSON, J, & HARRELSON, P V. The comparative abilities of White and Negro children *Comp Psychol Monog*, 1923, **5**

21 PETERSON, J. Comparisons of White and Negro children in the Rational Learning test Pub Sch Pub Co, 27th Yrbk Nat Soc. Stud Educ, 1928, Part I (P 333)

22 SCHMID, C F Social Saga of Two Cities Minneapolis Council of Soc Agencies, 1937 (Ch VII)

23 STRACHAN, L Distribution of intelligence quotients of twenty-two thousand primary school children *J. Educ Psychol.*, 1926, **16**. (P 169.)

24. SUNNE, D. Comparison of White and Negro children by the Terman and Yerkes-Bridges revisions of the Binet tests *J Comp Psychol*, 1925, **5** (P 209)

25 TERMAN, L. M., & MERRILL, M. Measuring Intelligence New York Houghton-Mifflin, 1937

Lubbock Field
Lubbock, Texas

SHORT ARTICLES AND NOTES

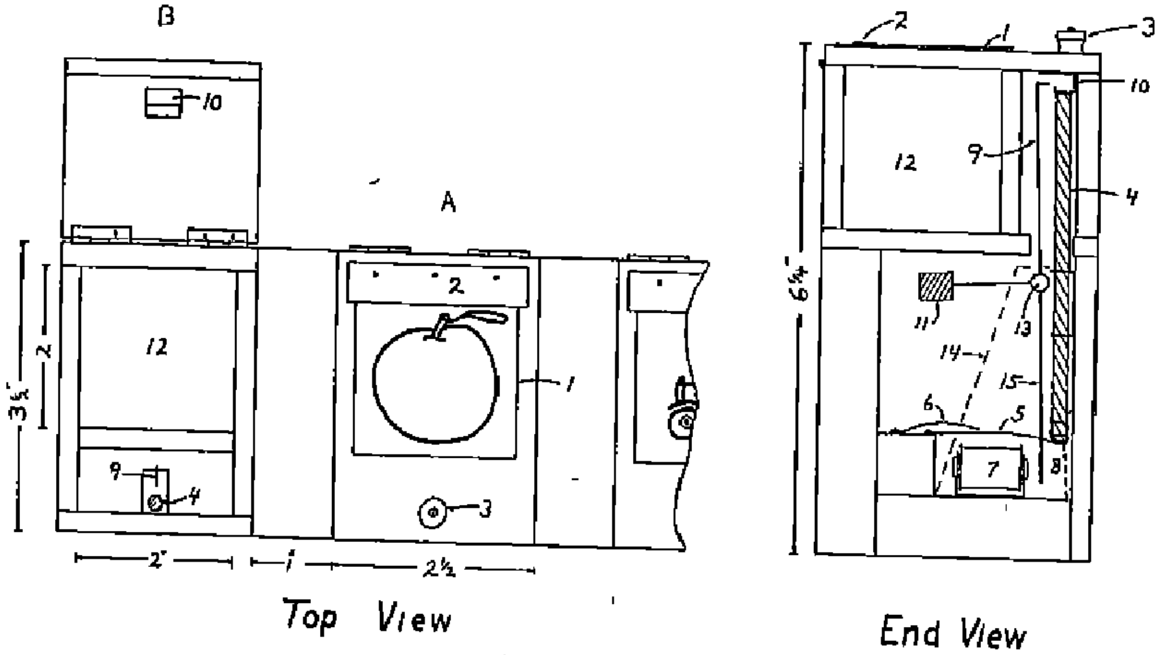
The Journal of Genetic Psychology, 1944, 65, 177-179

APPARATUS FOR THE STUDY OF DELAYED REACTION IN CHILDREN*

Institute of Child Welfare, University of California

Robert O. Scow and Margaret E. Robins

The apparatus herein described has been designed for the study of memory in young children. It presents a form of multiple choice situation consisting of 12 identical small boxes arranged in two banks of six boxes each. The



Top View

End View

FIGURE 1

Figure 1, End View, represents a lateral cross section of one of the six boxes constituting a bank of boxes. It shows the position of the stimulus card (1), the knob, for opening the hinged lid (manually) (3), the reward container (12), and the automatic locking device mentioned above. The essential parts and functions of this lock are as follows:

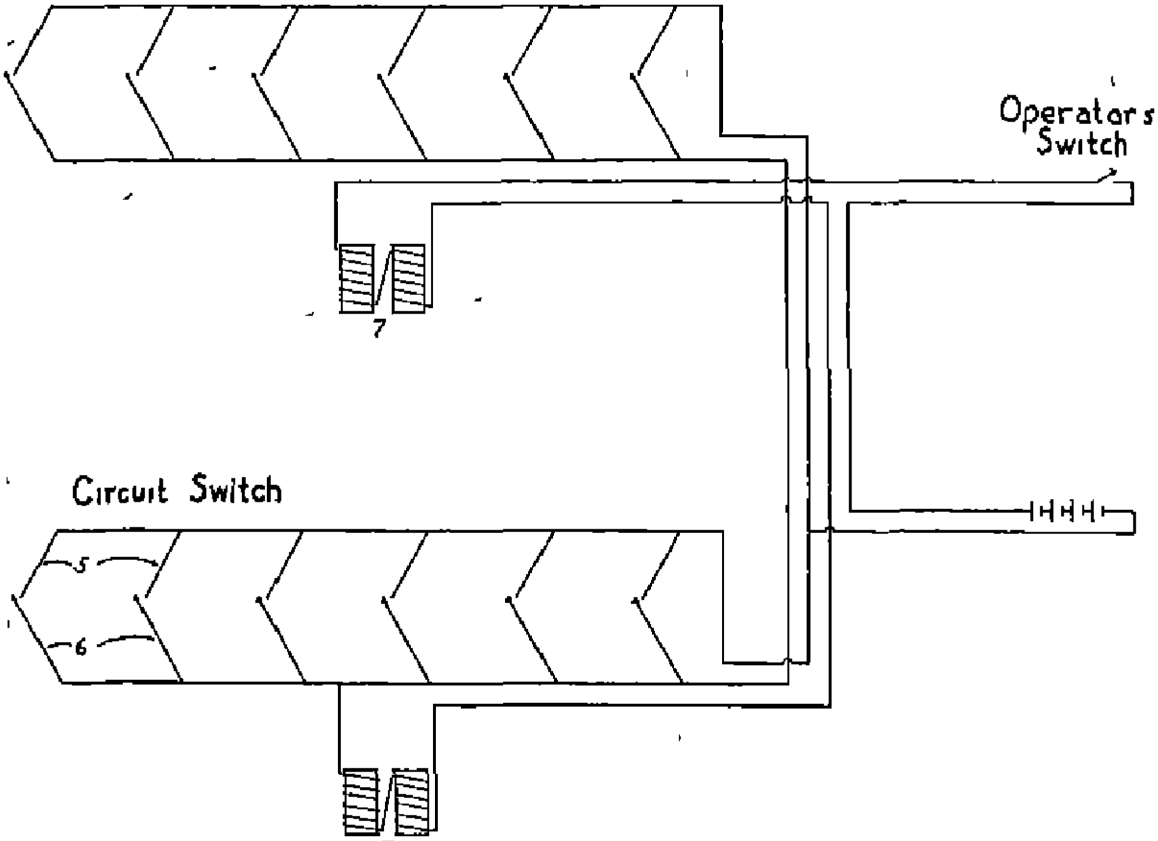
When the lid of the box is closed, as in the diagram, the hook (10), which is fastened to the lid, pushes a bar (4) downward, thereby separating contact springs

---

*Accepted for publication by Harold E. Jones of the Editorial Board, and received in the Editorial Office on June 12, 1942. Publication was delayed by the authors.

(5) and (6). These springs constitute a "circuit" switch Each box is equipped with a similar hook, bar, circuit switch and a latch (9). The latch in each box is attached to a common shaft (13) permitting a simultaneous movement of all six latches In the first box in each bank of six, the latch extends from the shaft downward, this extension is in the form of a rod (15), at the end of which an armature (8) is attached in juxtaposition to a small electromagnet (7) The opening of any one of the lids automatically closes the circuit switch below it

External features of the apparatus can be seen in Figure 1, Top View. Here one box (*B*) is shown with the lid open and another (*A*) with the lid closed



WIRING DIAGRAM

FIGURE 2

The schematic wiring diagram of Figure 2 shows the electromagnet, the power supply, and the switch controlled by the operator, connected in series, while the circuit switches (in the boxes) are connected in parallel It follows that, if one of the box lids in the bank is opened, the parallel circuit is closed at that point by the circuit switch controlled by that lid If the operator's switch is now closed (while the lid is still open) the circuit is complete, the electromagnet attracts the armature, and all the latches fly into the locked position This means that all the lids but one are now locked from the inside, and the opened one is prevented from being closed again by the position assumed by its own latch in relation to the hook

When the operator's switch is released, breaking the circuit, the boxes are automatically unlocked, and closure of the opened box made possible by the action of a counterweight (11, Figure 1), which returns the latches to the original "open" position

Dimensions for the apparatus are given in the diagrams, with the exception of total length, which measures 20 inches apparatus features a locking device which, upon the closing of the operator's switch, serves to fasten the lids of all closed boxes as soon as the lid of one of the boxes has been lifted This device restricts the subject to a single choice on any given trial, further, it allows additional time for the scoring of the response by preventing closure of the box which has been opened. The specific problem for which the apparatus has been employed concerns delayed recall of preschool subjects for meaningful and nonsense designs In investigation of this problem, the subject was tested for memory of the particular design previously associated with a reward Tests consisted of determination of the subject's ability to select this design from among a series of 12 designs. These tests were administered after certain predetermined intervals of time, placement of design and accompanying reward having been shifted during delay periods to control for the operation of positional cues However, the apparatus is suited for the investigation of a number of other problems and—with minor alterations—of other age groups The accompanying diagrams depict this apparatus in some detail (Figure 1).

The writers gratefully acknowledge suggestions made by Dr. H E. Jones, Director of the Institute of Child Welfare, Berkeley, California, and aid in preparation of this account given by Mr. Paul Lofgren, Psychology Department, University of California, Berkeley, California

Institute of Child Welfare
University of California
Berkeley, California

The Journal of Genetic Psychology, 1944, 65, 181-194

BOOKS

The *Journal of Genetic Psychology*, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than $2 per printed page and not more than $3 per printed page, but not more than $15.00 for a single review.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (*a*) Makes an important theoretical contribution, (*b*) consists largely of original experimental research, (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology, (*d*) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

(1942)

Sheldon, W. H. The Varieties of Temperament. New York: Harper, 1942. Pp. 520.

Klopfer, B., & Kelley, D. M. The Rorschach Technique. New York: World Book, 1942. Pp. 436.

Schilder, P. Goals and Desires of Man. New York: Columbia Univ. Press, 1942. Pp. 305.

Miller, J. G. Unconsciousness. New York: Wiley, 1942. Pp. 329.

Allport, G. W. The Use of Personal Documents in Psychological Science. New York: Social Science Research Council, 1942. Pp. 210.

Bender, I. E., Imus, H. A., Rothney, J. W. M., Kemple, C., & England, M. R. Motivation and Visual Factors. Hanover, N. H.: Dartmouth College, 1942. Pp. 369.

Boring, E. G. Sensation and Perception in the History of Experimental Psychology. New York: Appleton-Century, 1942. Pp. 644.

Groves, E. R., Groves, G. H., & Groves, C. Sex Fulfillment in Marriage. New York: Emerson, 1942. Pp. 319.

Watson, G. (*Ed.*) Civilian Morale. New York: Houghton Mifflin, 1942. Pp. 463.

Biber, B., Murphy, L. B., Woodcock, L. P., & Black, I. S. Child Life in School. New York: Dutton, 1942. Pp. 644.

Hollingworth, L. S. Children Above 180 IQ. New York: World Book, 1942. Pp. 332.

Klüver, H. (*Ed.*) Biological Symposia: Visual Mechanisms. Lancaster: Cattell Press, 1942. Pp. 322.

Landis, C., & Bolles, M. M. Personality and Sexuality of the Physically Handicapped Woman. New York: Hoeber, 1942. Pp. 171.

Lewis, T. Pain. New York: Macmillan, 1942. Pp. 192.

McGeoch, J A The Psychology of Human Learning New York, Longmans, Green, 1942 Pp 633

Schilder, P Mind Perception and Thought in Their Constructive Aspects New York, Columbia Univ Press, 1942, Pp 432.

Rees, H E A Psychology of Artistic Creation. New York Teach Coll, Columbia Univ., 1942 Pp 209

Seashore, C E Pioneering in Psychology Iowa City Univ Iowa Press, 1942

Rapaport, D Emotions and Memory Baltimore, Williams & Wilkins, 1942 Pp 282

(1943)

Gesell, A, & Ilg, F L Infant and Child in the Culture of Today New York Harper, 1943 Pp. 399.

Leeper, R W. Lewin's Topological and Vector Psychology Eugene Univ Oregon 1943 Pp 218.

Yerkes, R M. Chimpanzees A Laboratory Colony New Haven Yale Univ Press, 1943. Pp. 321

Young, P T. Emotion in Man and Animal New York Wiley, 1943 Pp 422

Masserman, J H Behavior and Neurosis Chicago: Univ Chicago Press, 1943 Pp 269.

May, M. A A Social Psychology of War and Peace New Haven Yale Univ. Press, 1943 Pp 281

Wolff, W The Expression of Personality New York Harper, 1943 Pp. 334

Estabrooks, G. H. Hypnotism. New York Dutton, 1943. Pp 249

Barker, R G, Kounin, J. S, & Wright, H F Child Behavior and Development New York McGraw-Hill, 1943 Pp 652

Newcomb, T M Personality and Social Change. New York: Dryden, 1943 Pp 225

McGraw, M. B. The Neuromuscular Maturation of the Human Infant. New York Columbia Univ Press, 1943 Pp 140

Child, I L. Italian or American? The Second Generation in Conflict. New Haven Yale Univ. Press, 1943 Pp. 208.

Griffith, C. R Principles of Systematic Psychology Urbana, Ill Univ Illinois Press, 1943 Pp. 718

Meier, N C Military Psychology. New York Harper, 1943 Pp. 395

Strong, E K, Jr Vocational Interests of Men and Women Stanford University Stanford Univ Press, 1943. Pp 746

Tomkins, S S (*Ed*) Contemporary Psychopathology Cambridge Harvard Univ Press, 1943. Pp 600

Hull, C L Principles of Behavior. New York Appleton-Century, 1943 Pp. 422

(1944)

Cantril, H Gauging Public Opinion Princeton Princeton Univ. Press, 1944 Pp 318

Hunt, J M (*Ed.*) Personality and the Behavior Disorders (2 vols.) New York Ronald Press, 1944 Pp 1242

CRITICAL REVIEWS OF RECENT BOOKS

The Journal of Genetic Psychology, 1944, 65, 183-191

(*Hollingworth, L S Children Above 180 IQ New York: World Book, 1942 Pp 332*)

Reviewed by W. Drayton Lewis

Dr. Leta S Hollingworth's major interest during the last two decades of her life was superior children, especially those who were unusually gifted. She was the moving spirit behind the two special classes for gifted children in the New York City schools and had intimate contact with many of these children Her first volume, *Gifted Children,* appeared in 1927 and she had planned to bring her material up to date Only three chapters of this projected book had been written at the time of her death in 1939. Dr. Harry L Hollingworth, her husband, has endeavored to preserve as much of her work as possible. The book contains the three chapters which she wrote, some case histories which she had written and which are supplemented by editorial notes, chapters written from material in her files, summaries written by the editor, while Part III is composed of papers which she had written during the last few years of her life

The material in this book is of sufficient value to justify its preservation in this form although it is to be sincerely regretted that Dr. Hollingworth could not have finished the book since much is lost that was her own personal information and which, of course, was not to be found in her files The book is an important contribution to the field in spite of the way in which it had to be written

Dr Hollingworth relates how her interest in superior children had its inception in 1916, which was about the time a suitable measuring device became available, that is, the Stanford revision of Binet's test. Throughout the remaining years of her life she collected material by a close study of many children of superior intelligence who came to her attention. A relatively large number of these children were studied by her over a period of years so that a large part of the material included in the book originated with her. The book, to a very large degree, may be considered a summary of her work and findings It is probably true that she had more direct contact with children obtaining *IQ*'s of 180 or above on the Stanford-Binet than any other individual has had. It is obvious that, this being the case, a great deal has been lost through the necessity of having so much of the material in the book prepared by an editor

Dr. Hollingworth was not of the opinion that a systematic search for these unusually able students is feasible. "*E*," who first came to her attention in 1916, was the first of these very brilliant children whom she studied and from that date she was continually seeking these rare individuals. She states that the search was conducted in a "desultory" manner and in "odd" moments from that date. The only occasion when the search was systematic was during the periods when children were being sought for the special classes at Public School 165, Manhattan, in 1922-23, and again in 1935-36 when another special class was set up at Public School 500, Manhattan. Her conclusion relative to the possibility of a systematic search for such individuals is: "In the course of long searching, I have learned that it is nearly useless to *look* for these children, because so few of them exist"

Valuable and significant material is reported to have been omitted out of deference to the well being of the children and in order to protect personal privacy Children of this ability read widely and are very likely to come across accounts of themselves This had to be considered in the preparation of the material included in this book It is to be regretted that much of value was lost because of this but one must recognize with the author the need of protecting these children

A persistent problem in the study of superiority is that of the concept of "genius" The term is not applied to those whose *IQ*'s are 180 or more, because no satisfactory or generally acceptable definition has yet been presented in the literature and because not all of these very superior youngsters have given real evidence of genius The viewpoint of this study is that it will be many years before we will know to what degree these children are "geniuses" It is also set forth as possible that no meaning for the word "genius" can be agreed upon in the end. The position taken is that, "All we know about the status of the subjects of the present study is that they test above 180 (*S-B*) and are thus more than +10 *PE* removed from mediocrity in general intelligence"

The concept of "genius," as it has appeared in the literature from the time of the Greeks to the present day, is discussed. The second chapter, which deals with early scientific studies of eminent adults, is sketchy and was probably to have been expanded. The bibliography constitutes the most valuable part of this section.

The choice of 180 *IQ* is admittedly arbitrary and was selected merely for the purpose of having a point at or above which very few children will score. It is pointed out that the frequency of such children is not known, they may be one, two, or three in a million. Several of the cases cited lived before the days of mental testing but are included because they obviously were quite unusual. It is also stated that John Stuart Mill, Leibnitz, Macaulay, Goethe and others probably belonged in this group. It has always seemed to the reviewer that studies such as this would be strengthened if less space were given to children who flashed on the horizon for only a moment and if more consideration were given to the early development of men, such as those named above, whose adult achievement have placed them in the genius group.

The remainder of the third chapter is devoted to studies of 19 children of tested *IQ*'s of 180 or more that have appeared in the literature. It is of interest to note that 12 of these are girls, 7 boys, whereas the 12 cases studied personally by Dr. Hollingworth included 8 boys and 4 girls, making a total of 16 girls and 15 boys. This might suggest that the sex distribution is about even among these very superior children. It must be pointed out, however, that in Dr. Hollingworth's studies of gifted children, as well as in that of Terman, there is an as yet unexplained male bias.

The next 12 chapters consist of 12 studies of children of *IQ* 180 or more who were followed by Dr. Hollingworth over a period of years. She had not written up many of these at the time of her death, so that a number of the chapters were written from her files by the editor. Thus, much has been lost. The studies of "*A*," "*B*," "*C*," and "*D*" end in 1923, 1925, 1925 and 1926 respectively and these are supplemented by brief notes by the editor. The study of "*E*" was culled by the editor from three studies of the child which Dr. Hollingworth had written. The remaining seven studies were written by the editor and some are quite brief. It is of interest to note that child "*I*" is the daughter of one of the male children reported in Terman's *Genetic Studies of Genius*. The reader is impressed by the fact that so many of the siblings of these children also have high *IQ*'s.

Significant findings of these individual studies are summarized by the editor in Chapters 16 and 17. Talking and reading, both of which involve the use and understanding of symbols, most clearly differentiate the records of these children from the norms. The reader is struck by the difficulty these children have in making adjustments in school, particularly in the elementary school These children are usually accelerated in school and very often troubled by those phases of school work which demand muscular or motor coordination. Motor development fails to keep pace with mental development, thus causing the children to have a good deal of difficulty in writing, handwork, and shopwork They are also handicapped in playground competition.

> Such children as are here presented constitute difficult educational problems from their entrance in school. The problems are not only those of the teachers and educational authorities, but they are chiefly, perhaps, the problems of the children themselves

The reader is also impressed by the fact that where the parents and teachers were understanding and sympathetic successful adjustments were usually made. It is striking to note how frequently parents and teachers do not recognize that these children possess any rare gifts. The adults do not know what to look for

The idea of creative ability or originality is usually associated with the concept of genius. Do these children possess this capacity or are they merely extremely rapid learners? Some of them seem to possess creative ability but it is impossible to say whether all do. The education and training of some of them has been such that it probably would have stifled such abilities had they been present It appears that in some cases there has been so much emphasis on rapid promotion that little time was left for creative activity. Dr Hollingworth did not believe that a socially significant answer relative to the creative ability of these children could be obtained until a study had been made of their adult achievements. The question as to what type of education is best adapted to the native creative ability of these children has yet to be answered.

The editor estimates that one-third of these highly intelligent children show notable signs of creativity, one-third give indication of creativity to "a moderate degree" whereas the others give no indications of marked constructive originality Dr. Hollingworth believed that the best educational program for these children would require that only a part of the child's time be devoted to the regular curricula, leaving time for the encouragement of individual initiative and enrichment. This was the program used with the special class at Speyer School in the late thirties where only half a day was given to the regular curricula. This program was based on "measurements" which show that a child of 140 *IQ* can master all the mental work provided in the elementary school, as established, in half the time allowed him, while a child of 170 *IQ* can do all the work required of him in one-fourth of the time he is compelled to spend in school

The last three chapters, which comprise Part III, are reprints of articles of Dr Hollingworth which appeared in various journals The subject matter here is the upper one or two per cent of the distribution rather than those of *IQ* 180 or above.

Chapters 18 and 19 deal with the adult status and personality ratings of superior children. An objection is raised to Terman's use of the term "genius" to denote children testing at or above 140 *IQ* (*S-B*) since this would classify as "geniuses" individuals who merely represent the upper fourth of all students being graduated from American colleges. It is stated that children with *IQ*'s of 180 or more rise above the generality of the population in adolescence and early maturity, that they surpass children who test around 140, 150, or 160 *IQ* (*S-B*) in honors and prizes won or in test score, and that it may be predicted with confidence that they will win prizes and honors for intellectual work. It is suggested that 180 *IQ* may be the point where the term "genius" begins to apply

As for personality ratings, it is stated that children testing from 135 to 190 *IQ* are "less neurotic, more self-sufficient, and less submissive, as a group, than are the populations with which they are comparable," that is, with "college students in general, or than adults of the mental caliber represented in the Bernreuter norms"

These superior children are not, as a group, socially annoying. They are, as a group, tall, heavy, strong, healthy, and fine looking but they are at a disadvantage in early years since they are apt to be accelerated in school. As they approach maturity this becomes less of a problem and "they reap the benefits of superior vitality, size and beauty" It is suggested that the "inferiority complexes" of gifted children arise out of their disadvantage physically with their schoolmates in early school years.

Problems of educational and vocational adjustment arise out of the superior child's varied skills. They learn so readily that they are apt to develop poor work habits and they have such varied abilities "that they may have difficulty in confining themselves to a reasonable number of enterprises"

One of the major problems of social adjustment for these brilliant children is that which Dr. Hollingworth calls "learning to suffer fools gladly" Not only children but adults do and say things that seem to them to be utterly absurd and foolish. Failure to learn this lesson often leads to "bitterness, disillusionment, and misanthropy"

Children of this very superior group are not, as a rule, unfriendly but they do have trouble finding companions who talk the same language and have the same interests. There are so few who are like minded. They tend to be incessant readers and this interest in reading also tends to isolate them These problems, however, tend to become really serious only when the child is so far above his group in intelligence that adjustment becomes almost impossible

One of the more interesting concepts developed by Dr. Hollingworth is that of "optimum intelligence." The idea behind this concept is that a certain amount of superiority in intelligence is an aid in adjustment but that a child can be so intelligent that adjustment is extremely difficult In other words, children of a moderate degree of acceleration mentally tend to make better adjustments than those whose intelligence quotients are extremely high The range of optimum intelligence is set between 125 and 155 *IQ* "Children and adolescents in this area are enough more intelligent than the average to win the confidence of large numbers of their fellows, which brings about leadership, and to manage their own lives with superior efficiency Moreover, there are enough of them to afford mutual esteem and understanding."

The children with *IQ* 170 or above are often unable to find congenial companions, they are "too intelligent to be understood by the general run of persons with whom they make contact" The children in the former group appear to find the world well suited to their development, the latter often do not "If a parent would want his child to enjoy 'every advantage' he could not do better than wish the child to be endowed with an *IQ* not lower than 130 or higher than 150." The reviewer believes that Dr. Hollingworth is correct in this concept, educational and social life being what it is

Dr. Hollingworth believed that two of the most important educational problems, if satisfactory personality adjustment is to be attained for these children, are (*a*) how to provide against alienation from contemporaries of both sexes and (*b*) how to prevent the negativism that results from continuous living under inefficient or unreasonable authority.

The problems of children of superior intelligence dealt with in Chapter 20 are essentially problems of the child rather than of society and arise from "the combination of immaturity and superiority." There is a great deal of repetition here of material already presented in early chapters but this is not surprising since these chapters are merely reprints of papers already published in journals with no attempt at integration. The specific problems discussed are those of (*a*) work, (*b*) adjustment to classmates, (*c*) play, (*d*) conformity, and (*e*) origin and destiny

The gifted child in the "lock-step" of the ordinary school has so little to do that he received daily practice in habits of idleness and day dreaming and the child of *IQ* 180 or above is apt "to regard school with indifference or positive distaste, for they find nothing interesting to do there." Two methods of meeting this problem are being tried: (*a*) special classes in populous centers, and (*b*) "a moderate degree of acceleration, combined with enrichment of the curriculum for the individual" Dr Hollingworth appears to have become quite committed to separate classes for very superior children wherever they are possible.

Superior children frequently have imaginary playmates, apparently because of their inability to adjust to schoolmates They are often unpopular in the play groups because "they want to organize play into a complicated pattern, with some remote and definite climax as a goal" Their language is not the language of their contemporaries "Literally, they could not understand each other." Girls of high intelligence are said to be, as a group, more competitive, aggressive, and active than girls are supposed to be These girls have problems which arise from their sex "Her problem is to adjust her ambition to a sense of sex inferiority without, on the one hand, losing her self-respect and self-determination, and, on the other, without becoming morbidly aggressive" This is not an easy adjustment.

Children of superior intelligence are, as a group, more easily disciplined than children in general but they are apt to be problems in the classroom. They do have ideas, lots of them, and the ideas will come out, they press for expression If there are several of these children in a classroom, they may all try to talk at the same time and confusion may result. "To hold his tongue, to listen quietly and respectfully to others, to speak according to some order of procedure, and to restrain disappointment at failure to be heard at all—these habits seem especially difficult for gifted children to form." Only gradually do these children learn self-government in this respect.

The gifted child's problems are usually those of childhood and immaturity With the passing of the years, the school group becomes more and more selective and the disparity between the bright child and his classmates is not so great. The nearer the child is to maturity the fewer and less acute his problems become.

Dr. Hollingworth sets forth, in the chapter on "The Elementary Schooling of Very Bright Children," four principles which she believes should be basic to any program of education for them (*a*) The education should afford them a rich background of ideas, in terms of which they may perceive the significant features of their own time (*b*) All subjects should be avoided which they will have occasion to encounter in high school and college in later years (*c*) The education of the best thinkers should be an education for initiative and originality. (*d*) They need to know the evolution of culture as it has affected common things

The program of training which was developed for the special class of gifted children at Speyer School in the late thirties is discussed in the chapter.

Five problems of general conduct which she believed have to be taken into consideration in developing a program for the emotional education of superior children, since they tend to beset these children in their early years and are subversive to the development of fine leadership, are

> (1) To find enough hard and interesting work at school; (2) to suffer fools gladly; (3) to keep from becoming negativistic toward authority, (4) to keep from becoming hermits, (5) to avoid the formation of habits of extreme chicanery

The final chapter is a paper read by Dr Hollingworth in 1939 and undoubtedly represents her final views on the education of children of superior intelligence. The two experiments in the education of the gifted which were conducted under her direction are described here Nine questions relative to the education of the gifted are listed "which call urgently for discussion as a joint responsibility of both elementary and secondary schools Some of these questions are discussed in this chapter. Reasons for selecting 13 as the age of transition to senior high school is presented. "Entering high school near the 13th birthday, a child saves time, and yet is not made subject to tensions which may result from trying to meet social and physical requirements for which he is too immature."

It was Dr Hollingworth's opinion that only those with *IQ*'s above 150 should have an enrichment program in high school.

> From observations of the progress of highly intelligent children tested at an early age, I offer the hypothesis that pupils of 130 to 150 *IQ* (*S-B*) have quite enough to do in the truly efficient pursuit of the college preparatory curriculum of the senior high schools, and do not

> need any enrichment of this curriculum as far as challenge to ability is concerned What these pupils need is merely freedom from the presence of great masses of classmates who are mentally unadapted to the college preparatory course, and the opportunity to work unhampered, in segregates groups, such as are now being formed in many secondary schools under the concept of *the honor school*

Dr Hollingworth, in the end, appears to have come to the conclusion that the concept "genius" should be reserved for those testing above *IQ* 170 or 180

> They are the ones who cannot only *conceive* thought in its abstract reaches, but who can *originate* new thoughts, new inventions, new patterns, and who can solve problems
>
> That point in the distribution of *IQ* where mental products suggestive of genius, as defined by lexicographers, begin to appear, seems to be as far above 140 *IQ* as 140 *IQ* is above average Somewhere between 170 and 180 *IQ* (*S-B*) we begin to see merging in early adulthood that "highly unusual power of invention or origination," that "original creative power, frequently working through the imagination," which is ordinarily called "genius."

Her challenge to educators is found in the following sentence

> The most interesting problem in education is to discover how these children, testing above 170 *IQ* (*S-B*), can and should be educated, to devise ways and means whereby these far deviates may get the full use of their abilities in school and society, especially when they have no money.

It is inevitable that the book is not well integrated and that it is sketchy in many places. It is indeed unfortunate that Dr Hollingworth was unable to give a final summary of her work She made a highly important contribution in her study of the traits and educational problems of these highly gifted children It seems logical to assume that these children are capable of making highly significant contributions to society and it is sincerely to be hoped that she may have stimulated enough interest in the educational possibilities of such superior individuals that some satisfactory solution of their problems will be achieved in the near future so that society will gain the full benefit of their abilities.

Department of Education
Murray State Teachers College
Murray, Kentucky

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i e, the most recently received books will be found at the end of the list)

Newcomb, T M Personality and Social Change New York, Dryden, 1943 Pp 225

Davidson, H. H Personality and Economic Background New York King's Crown, 1943. Pp 189

Yerkes, R M. Chimpanzees A Laboratory Colony New Haven Yale Univ Press, 1943 Pp 321

McGraw, M B The Neuromuscular Maturation of the Human Infant New York Columbia Univ. Press, 1943 Pp 139

Leeper, R W Lewin's Topological and Vector Psychology Eugene Univ Oregon, 1943. Pp. 218

Harrell, F H Effect of Added Thiamine on Learning New York Teach Coll, Columbia Univ, 1943 Pp 55

Breckenridge, N E, & Vincent, E L Child Development Philadelphia Saunders, 1943 Pp 592

Child, I. L Italian or American? The Second Generation in Conflict. New Haven Yale Univ Press, 1943 Pp 208

Bills, A G The Psychology of Efficiency. New York Harper, 1943. Pp 361

Howard, F T. Complexity of Mental Processes in Science Testing New York: Teach Coll, Columbia Univ, 1943 Pp 54

McNally, H. J The Readability of Certain Type Sizes and Forms in Sight-Saving Classes New York Teach. Coll, Columbia Univ, 1943 Pp 71

Bond, G. L, & Bond, E. Teaching the Child to Read New York Macmillan, 1943 Pp 356

Boring, E G, Van de Water, M, *et al* Psychology for the Fighting Man Washington' Infantry Journal, 1943 Pp 456.

Griffith, C. R Principles of Systematic Psychology Urbana, Ill Univ. Illinois Press, 1943 Pp 718

Tomkins, Silvan S (*Ed*) Contemporary Psychopathology Cambridge. Harvard Univ Press, 1943 Pp 600

Murphy, A E The Uses of Reason. New York Macmillan, 1943. Pp 346

Levy, D M Maternal Overprotection New York Columbia Univ Press, 1943 Pp 417

Strong, E K, Jr Vocational Interests of Men and Women Stanford University Stanford Univ Press, 1943. Pp 746

Graubard, M Man's Food. Its Rhyme or Reason New York Macmillan, 1943. Pp 213

Williams, C T These We Teach Minneapolis Univ Minnesota Press, 1943 Pp 188

Rivlin, H N, & Schueler, H. (*Eds*) Encyclopedia of Modern Education New York. Philosophical Library, 1943. Pp 902

Clayton, A S Emergent Mind and Education New York Columbia University, 1943 Pp 179.

Srinivas, M N Marriage and Family in Mysore Bombay New Book, 1942 Pp 219

Meier, N C Military Psychology New York Harper, 1943 Pp 395

Baxter, B, & Cassidy, R Group Experience New York' Harper, 1943 Pp 218.

Whiteway, H L Scientific Method and the Conditions of Social Intelligence New York Columbia University, 1943 Pp 188

Jessup, B E Relational Value Meanings Eugene, Oregon Univ Oregon Press, 1943. Pp 175.

Glass, B Genes and the Man. New York Columbia University, 1943 Pp 386

Jennings, H H Leadership and Isolation New York, Longmans, 1943 Pp. 240

Lee, A M., & Humphrey, N D Race Riot. New York Drydend, 1943 Pp 143

Hollingworth, H L Leta Stetter Hollingworth Lincoln, Neb Univ Nebraska Press, 1943 Pp. 204

Lowinger, A The Methodology of Pierre Duhem New York Columbia Univ Press, 1941. Pp. 184.

Taliaferro, W H, *Ed.* Medicine and the War Chicago Univ Chicago Press, 1944 Pp 193.

Hull, C L Principles of Behavior. New York. Appleton-Century, 1943 Pp 422

Nash, A S, The University and the Modern World New York Macmillan, 1944 Pp 312

Kasanin, J. S Language and Thought in Schizophrenia. Berkeley: Univ California Press, 1944 Pp. 133.

Bowley, A. H. Guiding the Normal Child New York: Philosophical Library, 1943 Pp 174.

Cantril, H Gauging Public Opinion. Princeton Princeton Univ Press, 1944 Pp 318

Flesch, R Marks of Readable Style. New York. Columbia University, 1943. Pp 69

Sargent, S S. The Basic Teachings of the Great Psychologists. New York. New Home Library, 1944. Pp 346.

Hunt, J. M, *Ed* Personality and the Behavior Disorders. (2 vols) New York Ronald Press, 1944 Pp 1242

Klein, D B Mental Hygiene. New York Holt, 1944 Pp 498

Salter, A What is Hypnotism? New York Smith, 1944 Pp 88.

Barnouw, A J, & Landheer, B. The Contribution of Holland to the Sciences New York: Querido, 1943. Pp 373

Moore, U, & Callahan, C. C. Law and Learning Theory New Haven. Yale Law Journal, 1943 Pp 136.

Ribble, M A. The Rights of Infants New York Columbia Univ. Press, 1943. Pp 118.

Parker, C. M Steel in Action. Lancaster, Pa.: Cattell Press, 1943 Pp. 221

Woodworth, R S, & Sheehan, M R First Course in Psychology. New York Holt, 1944 Pp 445

Thorndike, E L Man and His Works. Cambridge Harvard Univ Press, 1943 Pp 212

Winn, R B, *Ed* Encyclopedia of Child Guidance New York. Philosophical Library, 1943. Pp 456.

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publications shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price will be $7 per volume.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinco or halftone cuts.

(6). The publication charges will be as follows:

4 page signature	$18
8 " "	24
12 " "	36
16 " "	48
20 " "	60
24 " "	72
28 " "	84
32 " "	96
36 " "	108
40 " "	120

charges are tentative and will be adjusted as justified by the sub- list (especially by the effect of the war in China and in Europe).

The accumulated units will be mailed to subscribers quarterly: April, July, and October. The units may vary in size, to accom- publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

December, 1944
Volume 65, Second Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND
JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under

the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

DECEMBER, 1944

(Manuscripts are printed in the order of final acceptance)

IQ constancy on the revised Stanford-Binet from the preschool to the junior high school level . . . 197
By Katherine P. Bradway

Early individual differences in visual and motor behavior patterns: A comparative study of two normal infants by the method of cinemanalysis 219
By Louise Bates Ames

Generalization of the concept of middleness 227
By V. Graham, T. A. Jackson, L. Long, and L. Welch

Binet *IQ* changes of orphanage preschool children: A re-analysis . 239
By Beth L. Wellman and Edna Lee Pegram

(OVER)

Copyright 1944, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879

Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass. under the Act of March 3, 1879

A preliminary study of audition in the white rat 265
By Allen T. Dittmann

The inheritance of shyness in dogs 275
By Frederick C. Thorne, M.D.

Infant learning as a developmental index 281
By John J. B. Morgan and Sarah S. Morgan

Age differences in religious beliefs and problems during adolescence 291
By Raymond G. Kuhlen and Martha Arnold

SHORT ARTICLES AND NOTES 301

An apparent case of neurotic influence on somatic function (Martin L. Reymert) 301

BOOKS 301

CRITICAL REVIEWS OF RECENT BOOKS 307

A. Gesell and F. L. Ilg, *Infant and Child in the Culture of Today* (Reviewed by Dorothy Tilden Spoerl) 307

BOOKS RECENTLY RECEIVED 315

The Journal of Genetic Psychology, 1944, 65, 197-217

IQ CONSTANCY ON THE REVISED STANFORD-BINET FROM THE PRESCHOOL TO THE JUNIOR HIGH SCHOOL LEVEL*

Department of Psychology, Stanford University

KATHERINE P. BRADWAY[1]

A. INTRODUCTION

Intelligence tests are used not only to measure present intelligence, but also to predict future intellectual functioning. The clinical psychologist's recommendations for a child, and the research investigator's studies of factors affecting intellectual development are both based on the assumption that the rate of mental growth, as indicated by the *IQ*, tends to remain constant. The development of mental tests has been accompanied by widespread investigations to determine the degree of this constancy. Children have been examined at various age levels and re-examined at later ages. Differences between *IQ*'s from the initial and retest examinations have then been submitted to statistical treatment in order to secure an index, usually a correlation coefficient, of *IQ* constancy. The large number of references listed in each of the several reviews of the literature on this problem (4, 14, 31, 35, 37, 40) is evidence of the extensive interest which it has aroused. Some of the factors which have been found to be important in their effect on *IQ* constancy however, were not controlled in most of the early studies. Two of these are the age of the child at the time of the initial examination and the length of the interval between the two examinations (1, 25, 41). The younger the child at the initial examination, and the longer the time interval between examinations, the less constant is the *IQ*. Only a minor number of investigators (11, 13, 27, 28) have concluded from their data that there was little or no relation between *IQ* constancy and initial age or interval.

---

*Received in the Editorial Office on January 9, 1943.

[1] This study was the preliminary part of an investigation reported in a dissertation entitled "An Experimental Study of Factors Associated with Stanford-Binet *IQ* Changes from the Preschool to the Junior High School Level" which was submitted to the Graduate School of Stanford University in 1942 in partial fulfillment of the requirements for the Ph.D. degree. The writer wishes to acknowledge her indebtedness to Dr. Maud A. Merrill for her aid throughout the study.

The majority of the mental growth studies have been made at the school level Greater availability of school children is partly responsible for this, but the later development of preschool tests as compared with tests for school-age children has been the major factor The 1916 Stanford-Binet, the test most frequently employed in mental growth studies up to 1937, was of limited use below the five-year level. The eventual development of preschool and infant tests, however, permitted the measurement of intelligence at younger ages and stimulated investigations of the predictive value of preschool *IQ*'s. If the age at the time of the retest was above four years, the Stanford-Binet was generally used for the re-examination The problem at the preschool level thus became one of determining the extent to which the results of a test given prior to school entrance predicted later Stanford-Binet *IQ*'s. A difficulty arises, however, in comparing *IQ*'s from different tests Differences in the populations on which the tests have been standardized tend to disturb the comparability of the *IQ*'s. The appearance of the Revised Stanford-Binet Scale in 1937 made available a reliable and carefully standardized measure of intelligence with a test-age range from 18 months to the superior adult level Studies of mental growth from the preschool to the adult level can now use the same test throughout the entire age period.

The present investigation is a 10-year follow-up study of children who had been examined between the ages of two and six years with the Revised Stanford-Binet Scale during the standardization of this scale. The data obtained from the follow-up study furnish the basis for an appraisal of the predictive value of Stanford-Binet *IQ*'s obtained at the earliest ages at which the revised scale is applicable The re-examination was made when the children were 12 to 16 years old Since mental growth is nearly or entirely completed at this age, the results give information regarding the extent to which the earliest obtainable Stanford-Binet *IQ*'s prognosticate adult *IQ*'s.

B. SURVEY OF THE LITERATURE

It is not within the scope of the present report to attempt to summarize the results of all of the many studies on *IQ* constancy There have appeared, however, reports of a few studies which include results on subjects who were initially examined at the preschool level—that is, between two and six years. These results are particularly pertinent to the present investigation and will be reviewed briefly Results of initial examinations made prior to the age of two years or subsequent to the age of six years will not be included. Studies in which the interval between examinations was one year or less will be omitted since the problem in such studies is primarily one of *IQ* reliability rather than *IQ* constancy Those studies in which a wide age range and a wide interval range were not broken down into smaller units will also be omitted since the results of such studies cannot be compared satisfactorily with those of studies such as the present one in which there were a small initial age range and a small interval range. The results will be summarized in terms of the correlation coefficient and, when it has been reported or can be readily calculated, the mean *IQ* change

Anderson (2) compared the results on the Gesell administered to 91 children at the age of two years with their *IQ*'s on the 1916 Stanford-Binet at the age of five years The mean *IQ* at five years was 116 with a range from 83 to 149. The standard deviation is not reported, but the range indicates a curtailment at the lower end of the *IQ* distribution and a consequent decreased variability Anderson takes account of this in his suggestion that the obtained correlations probably represent the minimum expression of the true relationship The correlation between Gesell scores at two years and Stanford-Binet *IQ*'s at five years was 45. The addition of the applicable items from the Merrill-Palmer and the 1916 Stanford-Binet to the Gesell items at two years—resulting in a total of 183 items—raised the correlation to .52 Nine subjects were added to the original 91 for this latter correlation. Scores on a scale made up of the 46 most diagnostic items of the total 183 correlated with five-year Stanford-Binet *IQ*'s to the extent of .55. The group on which this correlation was based included the validating group.

Bayley (3, 5, 6, 7), and Bayley and Jones (8), have reported data from the first Berkeley growth study on the constancy of the *IQ* over the age range with which the present study is concerned Forty-eight of the original group of 61 children were examined from infancy through nine years The California Preschool Scales were used from 18 months through the five-year level, the 1916 Stanford-Binet was used at ages six and seven years, the Stanford-Binet, Form *L*, was used at ages eight and nine. That these children represented a selected sample is indicated by the fact that their mean *IQ* at the age of nine years was 129 The standard deviation of the distribution of *IQ*'s at this age was 22, however, which indicates that the group was a heterogeneous one[2] The correlations which were found between initial mental test point scores at the preschool level and sigma scores derived from *IQ*'s at nine years were: 52 for two years, 57 for two and a half years, .59 for three years, .57 for three and a half years, 74 for four years, .85 for four and a half years, .82 for five years. A subscore for a group of items selected from the California Preschool Scale on the basis of their contribution to the total mental score failed to produce a higher correlation with eight-year Stanford-Binet *IQ* than did the total score. Moreover, a "developmental score" based on the mental and motor test scores combined had no better predictive value than did the mental test score alone.

---

[2] The standard deviations of Form *L* *IQ*'s at the several age levels in the standard-

Cunningham (10) found a test-retest correlation of 24 for 20 subjects first examined with the Kuhlmann-Binet at two years and re-examined with the 1916 Stanford-Binet at a median age of eight years (range 7-0 to 8-7) No measures of central tendency or dispersion of *IQ*'s are reported for this sample group, but the group of 37 subjects from which it was drawn had a range of *IQ* from 88 to 148 at the age of eight years. This *IQ* range is similar to that reported in the Anderson study and likewise indicates a curtailment of the distribution at the lower end. There is a corresponding probability that the true relationship is closer than is represented by the obtained correlation of .24

De Forest (11) administered the revised Stanford-Binet to 170 children who had previously taken the Merrill-Palmer tests. Forty-four of these children had also taken the Minnesota Preschool tests. Age at the time of the initial examination ranged from 24 to 63 months. The interval between tests ranged from one month to 88 months. The mean Stanford-Binet *IQ* for the total group was 113. The only index of *IQ* variability which is reported is *IQ* range This was 58 to 160 for the total group, which suggests a normal variability. De Forest found a correlation of .55 between Merrill-Palmer and Binet *IQ*'s for the total 170 children and a correlation of .56 between Minnesota and Binet *IQ*'s for the 44 children initially examined with the Minnesota tests. An average of the Minnesota and Merrill-Palmer *IQ*'s for this latter group correlated with later Stanford-Binet *IQ* to the extent of .61. The mean change from the Merrill-Palmer *IQ* to the Stanford-Binet *IQ* was 12 points The mean change for the Minnesota *IQ*'s is not indicated.

---

ization sample ranged from 12 to 21 points. Terman and Merrill have suggested that 16 points may be accepted as approximately the representative value of the standard deviation of Revised Stanford-Binet *IQ*'s for an unselected population (39, p 40)

De Forest made a more extensive analysis of the Merrill-Palmer and Stanford-Binet results for the total group with regard to the influence of test interval and initial age on *IQ* constancy This analysis failed to reveal a relation between length of interval and constancy correlations. The correlation for the sample which was re-examined after an interval of seven months or less was 54; the correlation for the sample which was re-examined after an interval of more than seven months was .58 This is not in keeping with the majority of results reported by other investigators An even more unusual finding was that as initial age increased, prognostic accuracy for later Stanford-Binet *IQ*'s decreased For 77 subjects who were first examined at 24 to 42 months the correlation was .71, for 93 subjects initially examined at 43 to 63 months the correlation was .45

Driscoll (12) analyzed the test-retest results of a group of children who were examined with either the Kuhlmann-Binet or the Merrill-Palmer at the age of two to four years and re-examined with the 1916 Stanford-Binet at the age of five years or older. The upper age limit at the time of the second examination was not indicated in the report of the study That the children were a superior group is indicated by their mean Stanford-Binet *IQ* which was approximately 122 Normal *IQ* variability is indicated, however, by the standard deviation of the distribution of *IQ*'s which was about 16 points.[3] The correlations between initial and retest *IQ*'s were as follows 66 for 39 children initially examined with the Kuhlmann-Binet at the age of two to three years, 61 for 21 children initially examined with the Merrill-Palmer at the age of two to three years, .71 for 54 children initially examined with the Kuhlmann-Binet at the age of three to four years, .41 for 39 children initially examined with the Merrill-Palmer at the age of three to four years

Hallowell (23) has recently reported the results of the continuation of a study, the original results of which were reported in 1932 (22) Follow-up studies were made of 250 children who were first examined when they were less than four years old The final examination was made when the children were between 5 and 13 years old The median retest age was six years. The 1916 Stanford-Binet *IQ*'s at the time of the final examination ranged from 42 to 151 with a median of 96 Hallowell did not limit her examination at the preschool level to the administration of a single scale. Instead she administered several tests The results of these tests coupled with clinical observations furnished the basis for deriving a developmental quotient The predictive value of these *DQ*'s was found to be higher than that which had been reported for single scale scores at these age levels. Correlations between *DQ*'s at the preschool level and retest Stanford-Binet *IQ*'s were: .80 for 150 children initially examined at the age of 24 to 35 months, 84 for 173 children initially examined at the age of 36 to 47 months Mean changes in quotients for these age levels were 10 and 8 points respectively.

---

[3] The representative standard deviation of 1916 Stanford-Binet *IQ*'s may be assumed to be between 12 and 16 points In the standardization study it was found that the range of the middle 50 per cent of *IQ*'s was 16 points (38, p. 40), which corresponds to a standard deviation of 12 points In the Scottish survey of nearly 1000 unselected children the standard deviation of the *IQ* was 16 points (32, p 645).

Honzik (24, 25) has reported results from the Berkeley Guidance study[4] in which 250 children were first examined at the age of 21 months and were re-examined at regular intervals The mental tests used were the California Preschool Scales I and II between the ages of 21 months and five years, and the 1916 Stanford-Binet at six and seven years. The standard deviation of six- and seven-year *IQ*'s was 13 points This value is within the range of normal variability. The mean *IQ* of this sample at seven years was 118 Correlations between initial sigma scores on Forms I and II of the California Preschool Scale and sigma scores converted from seven-year Stanford-Binet *IQ*'s were respectively. .30 and 26 for 21 months, .57 and .54 for three years, 59 and .58 for three and a half years, .59 and .58 for four years, 69 and 76 for five years. The number of cases on which these several correlations were based ranged from 73 to 181 The mean difference between the initial sigma score and the retest sigma score was roughly equivalent to 12 *IQ* points for the 21-month level, 9 *IQ* points for the three- and four-year levels, and 6 points for the five-year level[5]

Katz (26) has reported results on the constancy of the 1916 Stanford-Binet *IQ*'s obtained in the course of a longitudinal study being made at the Brush Foundation. There were 268 children who were initially examined with the 1916 Stanford-Binet at the age of three years and re-examined with the same scale at least three times up to the age of five years The mean *IQ* at the time of the five-year examination was 117 The standard deviation at this age was 13 points which indicates that the variability of the group with regard to *IQ* was within normal limits. The correlation between three-year and five-year *IQ*'s was .62 The mean change in *IQ* over the two-year period was 12 points.

---

[4] Results from this study have also been reported by Macfarlane (29)

[5] Honzik reported mean changes in terms of sigma scores The values presented here were calculated by multiplying the changes in sigma scores by the standard deviation.

In mental growth studies reported by Gesell, by Nelson and Richards, and by Wallin tests were administered during the preschool period and were followed by retests at later ages. The results were not treated statistically, however, and consequently there are no indices of the degree of *IQ* constancy Gesell and his colleagues (16, 18, 19) have recently reported the results of a growth study of 30 children who were initially examined at the Yale Clinic as infants. These children were in their 'teens or older at the time of the latest examination The 1916 Stanford-Binet was used for children of school age. Prior to that age developmental quotients were arrived at by relating the child's chronological age to a general maturity level which was clinically estimated on the basis of the child's previous history coupled with the current examination in which the Yale norms of behavior maturity were used. Gesell presented the results in the form of individual growth curves supplemented by interpretative statements. As was the case in his 1925 study (17), Gesell was impressed with the constancy of the mental growth patterns revealed in these children. "In no instance did the course of growth prove whimsical or erratic. The behavior biographies give clear evidence of a high degree of latent predictability, even in infancy" (18, p. 149).

A growth study at the Fels Research Institute has provided data on mental growth (33, 34, 36) At age 18 months the Gesell schedules and the Merrill-Palmer Scale were used, at ages 24 and 30 months the Merrill-Palmer was used, at three years and thereafter the 1916 Stanford-Binet and, more recently, the revised Stanford-Binet have been used Nelson and Richards (33 and 34) of the Fels Institute have reported correlations between initial tests and retests up to the age of three years, but no correlations for the later ages have as yet been reported. In a paper presented at the annual meeting of the American Psychological Association in 1940 Richards (36) presented several individual growth curves which showed the course of development from the age of about three years to the age of seven years or older. At that time Richards emphasized the need for directing attention to the analysis of individual cases

Wallin (44) made 27 semi-annual examinations of two sisters from the time they were three years old until they were fifteen-and-a-half years old The same scale, the 1916 Stanford-Binet, was used throughout and was administered always by the same person Considerable variation between successive *IQ*'s was found, and Wallin cautioned against the use of a single *IQ* for prediction. Between the ages of 3 and 14 the *IQ*'s for one subject fluctuated between 106 and 123, and those for the other subject between 105 and 124.

From this brief review it is evident that the results from studies made of the predictive value of preschool measures or estimates of intelligence are diverse The correlations between initial and retest *IQ*'s range between 24 and .85, and the reported mean *IQ* changes range between 6 and 14 points Table 1 has been prepared to facilitate comparison of the several correlations reported. It will be noted that the *r*'s in the table tend to decrease from left to right and increase from top to bottom, reflecting the influences of increase in interval and increase in initial age Further comparison necessitates referring to the summary of each study in order to note the tests used and the dispersion of *IQ*'s of the groups of subjects on which the correlations were based It may be observed in this connection that in each study in which a correlation was reported, with the exception of the studies of Anderson and Cunningham, the *IQ* variability was apparently within normal limits.

TABLE 1

SUMMARY OF *IQ* CONSTANCY CORRELATIONS REPORTED FOR PRESCHOOL CHILDREN

Age in years at first test	Age in years at retest 5	5 plus	6*	7	8	9	1 to 88 mos later
2	.45 52			30† .26†	24	.52	
2 5		.66 .61	80			.57	
2 to 3.5							.71
2 to 5							55 .56
3	62			57 54		59	
3.5		71 .41	.84	59 .58		57	
3 5 to 5							45
4				59 .58		74	
4 5						.85	
5				69 76		82	

*Median of age range 5 to 13

†Initial examination three months before second birthday.

C. SUBJECTS AND PROCEDURE

The subjects for the present investigation had been examined in California 10 years previously in connection with the standardization of the two forms of the Revised Stanford-Binet Scale. The California communities which had served as sources for subjects for the standardization were representative of the total standardization group in that they included one large city, one small urban community, and several rural districts. In the age range chosen for the present study—two through five and a half years—there were 213 children examined for the standardization in these California communities Although a number of the children had moved during the 10-year interim, 152 of the original subjects, or 71 per cent of the original group of 213, were located.

In a follow-up mental growth study of this type there are several considerations which are important in the selection of subjects. First, it is important that there be no question as to the validity of the initial examination. There is no necessity for restating the careful conditions under which the examinations for the standardization of the Revised Stanford-Binet Scale were administered and scored (39) Since the present study was concerned with individual *IQ* changes, all subjects who had failed *more than one item* at the two-year, or lowest, age level on both forms of the initial Stanford-Binet examination were eliminated from the sample This precaution was taken in order to avoid the possibility of including any subjects whose initial *IQ*'s had been artificially raised. It meant the loss of 13 subjects at the 2- and 2½-year levels and resulted in a rise of the mean *IQ*'s at those levels since the most retarded children were necessarily eliminated.

Establishment of the accuracy of all birth dates used is a second requirement in a mental growth study. Although this was given careful consideration in the standardization study, additional steps were taken to avoid including any subject about whose birth date there was doubt. In every case the birth date as given in the original lists was checked with the birth date indicated in the school records. Records of birth certificates were available in 30 per cent of the cases. In only one case was there a doubtful birth date which could not be verified. This case was eliminated from the study.

As a third precautionary measure, data were obtained as to whether any of the subjects had been examined with the Stanford-Binet Scale since the initial examination. This was done to guard against possible practice effects Ten of the subjects had been examined with the 1916 Stanford-Binet Scale *more than four years previously*. There seemed to be no reason for excluding these subjects. Only one subject had been examined with the Revised Stanford-Binet Scale since the initial examination Since this examination had been made eight months previously, and since the child denied recognizing the test items, this subject was also included in the study

The elimination of the 13 subjects who had failed *more than one item* at the lowest age level on both forms of the Stanford-Binet Scale in the initial examination and the elimination of the subject whose birth date could not be established decreased the total number of subjects to 138 These 138 subjects formed, according to the age at which they had been initially examined, two groups the 2- and 3-year group consisted of 52 children who had been initially examined at 2, 2½, 3, or 3½ years of age, the 4- and 5-year group consisted of 86 children who had been initially examined at 4, 4½, 5, or 5½ years. The two sexes were distributed about equally in the older group (45 boys and 41 girls) In the younger group there were about 75 per cent more girls than boys (19 boys and 33 girls). No attempt was made to treat the sexes separately.

Tables 2 and 3 are presented to show the extent to which the retest groups

TABLE 2

DISTRIBUTION OF INITIAL *IQ*'S FOR THE RETEST AND STANDARDIZATION SUBJECTS

	Aged 2 and 3 years				Aged 4 and 5 years			
	Retest group		Standardization group		Retest group		Standardization group	
IQ	No	%	No	%	No.	%	No	%
150-159	2	4	3	1				
140-149	1	2	5	2			1	3
130-139	9	17	12	4	3	3	9	3
120-129	10	19	50	17	9	10	25	7
110-119	11	21	67	23	19	22	75	22
100-109	10	19	64	22	26	30	99	30
90-99	5	10	45	15	14	16	82	24
80-89	4	8	29	10	12	14	35	10
70-79			14	5	3	3	9	3
60-69			2	6			1	.3
50-59			2	6			1	.3
40-49			1	3			1	3
30-39							1	.3
Total	52	100	294	100 5	86	98	339	100.5
Mean	116 3		107 3		104 9		103 7	
SD	17 1		17 7		14 1		14 3	

were representative of the standardization group The distributions of initial *IQ*'s (composite of Form *L* and Form *M*) are compared in Table 2, the distributions of paternal occupations (where known) classified according to the Minnesota Scale for Occupational Classification (21) are compared in Table 3. It will be noted that the distributions of *IQ*'s and of paternal occupations for the retest 4- and 5-year group are essentially similar to the standardization 4- and 5-year group. These data establish the representative character of the sample 4- and 5-year group so far as these two factors are concerned. In comparison with the standardization group, the retest 2- and 3-year group shows a shift toward higher *IQ*'s. This selection in the direction of *IQ* superiority is a result of the elimination from the retest group of the younger children who failed more than one item at the lowest age level on both forms of the initial examination. The paternal occupational distributions, however, for the retest and standardization groups at this age level are similar.

TABLE 3

DISTRIBUTION OF PATERNAL OCCUPATIONS FOR THE RETEST AND STANDARDIZATION SUBJECTS

Occupational class*	Aged 2 and 3 years				Aged 4 and 5 years			
	Retest group		Standardization group		Retest group		Standardization group	
	No	%	No	%	No	%	No	%
1	2	4	21	7	5	6	15	4
2	7	14	27	9	4	5	23	7
3	11	22	79	27	19	24	96	28
4	4	8	27	9	3	4	32	9
5	19	39	99	34	32	41	125	37
6	5	10	32	11	11	14	27	8
7	1	2	9	3	4	5	21	6
Total	49	99	294	100	78	99	339	99

*Minnesota Scale for Occupational Classification

Form *L* of the Revised Stanford-Binet Scale was administered to the retest group of 138 subjects during the first three months of 1941. All of the examinations were made by the writer without knowledge of the previously obtained results. To insure accuracy of scoring every instance of doubtful scoring was submitted to Dr. Merrill for final decision.

D RESULTS

At the time of the initial examination the two forms of the Revised Stanford-Binet Scale were administered to the subjects with an interval of from one to a few days between testings. Form *L* was administered first in about half of the cases, and Form *M* was administered first in the remaining half of the cases. In order to allow for practice effects an appropriate correction was applied to the *IQ* of the second test (39, p. 43). The *IQ* from the first administered test at the time of the initial examination is herein referred to as the "initial first-test *IQ*"; the *IQ* from the second administered test at the time of the initial examination is herein referred to as the "initial second-test *IQ*." Form *L* and Form *M* were administered as the first test about equally often for this sample. The composite *IQ* which is referred to in this study was the average of the *IQ*'s from the two forms of the test after the correction was applied to the second-test *IQ*.

Several comparisons may be made between the initial *IQ*'s and the retest *IQ*'s on Form *L* which was administered 10 years later.[6] These comparisons for the two age groups are indicated in Table 4. The first two columns for each age group present the means and the standard deviations for initially obtained *IQ*'s and retest *IQ*'s. The last column for each age group indicates the Pearson product-moment correlations, with their probable errors, between initially obtained *IQ*'s and retest *IQ*'s.

TABLE 4

MEANS, STANDARD DEVIATIONS, AND CORRELATIONS FOR INITIAL AND RETEST *IQ*'S

	Aged 2 and 3 (*N* 52)			Aged 4 and 5 (*N* 86)		
	Mean	*SD*	*r* with retest *IQ*	Mean	*SD*	*r* with retest *IQ*
Initial Form *L IQ*	115.7	17.0	.58± .06	105.9	14.9	.67± .04
Initial Form *M IQ*	116.3	18.1	.67± .05	105.4	13.3	.63± .04
Initial Composite *IQ*	116.4	16.9	.66±.05	105.1	13.9	.67± .04
Initial First Test *IQ*	116.0*	17.1	.65± .05	105.8†	13.7	.66± .04
Initial Second Test *IQ*	116.1**	18.9	.62±.06	105.1†	14.6	.64± .04
Retest Form *L IQ*	114.5	18.4		108.4	16.5	

*28 Form *L* and 24 Form *M*

**24 Form *L* and 28 Form *M*

†Forms *L* and *M* divided equally (43 of each)

It will be noted that the correlation for the composite *IQ*'s at each of the two age levels is reached or exceeded by the correlation for the *IQ*'s on one of the two single forms. The correlation between Form *M IQ*'s at the 2- and 3-year level and retest Form *L IQ*'s is .67, which is .01 higher than the correlation between composite *IQ*'s at the 2- and 3-year level and retest Form *L IQ*'s. The correlation between Form *L IQ*'s at the 4- and 5-year level and retest Form *L IQ*'s is also .67 and is identical with the correlation between composite *IQ*'s at the 4- and 5-year level and retest Form *L IQ*'s. The correlation of prediction is higher for Form *M* than for Form *L* initially administered at ages 2 and 3, but the correlation of prediction for ages 4 and 5 is higher for Form *L* than for Form *M*.

In contrast to the results of Goodenough (20) and Updegraff (41) the findings of the present study do not suggest that the *IQ*'s of a first examination administered to young children are less reliable than are *IQ*'s obtained on a second examination administered shortly after the first one. As indicated in Table 4 the correlations between the initial first-test *IQ*'s and retest *IQ*'s for the two age levels are 65 and 66 respectively, whereas the correlations between the initial second-test *IQ*'s and retest *IQ*'s are .62 and 64 respectively. Thus the *IQ*'s obtained the first time the subjects of this study were examined were just as reliable indices of future *IQ*'s as were *IQ*'s obtained at the second examination administered within a few days of the first one

---

[6] The interval between the two examinations was not 10 years in all cases. It ranged between 8.5 and 11.0 years. The interval for 55 per cent of the subjects was between 9.5 and 10.5 years.

In general the correlations reported in Table 4 are similar to one another With the exception of one correlation, they range between .62 and 67. None of the differences between the correlations presented in the table is statistically significant The largest difference within an age group is that between the correlations of 58 for Form *L* and .67 for Form *M* at the 2- and 3-year level The difference between the *z*'s corresponding to these correlations divided by the standard error of the difference was found to be .94. This corresponds to a probability of .347 which indicates that the observed difference between the two correlations would arise 34 or 35 times in 100 solely through errors of sampling (42). Hence the present data do not justify the conclusion that at age two and three years Form *M* is a better test than Form *L* for predicting future intelligence

The question arose as to whether the obtained differences in correlations for the two forms at the two age levels (superiority for Form *M* at age 2 and 3 and superiority for Form *L* at age 4 and 5) were a reflection of differences between the predictive values of different age levels on the scale So that this might be investigated the subjects were grouped according to mental age rather than chronological age, and correlations for the two forms at two mental age levels were calculated. At both mental age levels the direction of the difference between the correlations was reversed. For the 54 subjects in the 2-and 3-year mental age group the test-retest correlations were 61± 06 for Form *L* and 60± 06 for Form *M*. For the 84 subjects in the 3- and 4-year mental age group the correlations were .66±.04 for Form *L* and .71±.04 for Form *M* The differences for both groups are too small to be statistically significant These reversals merely add weight to the conclusion that the differences found between the correlations for Form *L* and Form *M* for the chronological age groups in this sample may be due to chance and are not indicative of true differences.

Comparison between the results for the 2- and 3-year group and those for the 4- and 5-year group fails to reveal significant differences. In these results there is little basis for contending that Stanford-Binet *IQ*'s at ages 4 and 5 are more valid indices of future *IQ*'s than are Stanford-Binet *IQ*'s at ages 2 and 3. Correlations for the composite *IQ*'s at the two age levels are almost identical (.66 and .67). However, the somewhat greater homogeneity of the older group makes the slight superiority in correlations for this group more significant. Moreover, the results in Table 5 suggest that *IQ*'s at the higher age level have the greater predictive value. This table will be discussed presently.

The correlations in Table 4 may be compared with those which have been reported by other investigators. It will be recalled from the survey of the literature which was made earlier in this report that the correlations found in previous studies represent the extent to which preschool *IQ*'s on some test other than the Revised Stanford-Binet Scale predict later *IQ*'s on either the 1916 or revised Stanford-Binet scales. The present correlations for the 2- and 3-year group for Form *M* *IQ*'s and composite *IQ*'s (.67 and .66 respectively) are higher than correlations previously reported for these ages when the interval between tests was more than five years. The longest interval for which constancy correlations for these ages have been reported was seven years, this was the longest interval obtaining in Bayley's retests of children initially examined at the age of two years on the California Preschool Scale. Bayley found a correlation of .52 between initial test scores and *IQ*'s obtained seven years later. The correlations for children initially examined at 2½, 3, and 3½ years and retested from 6½ to 5½ years later were respectively .57, .59, and .57. Honzik's correlations for children initially tested at 3 and 3½ years on the California Preschool Scale and retested 4 to 3½ years later on the Stanford-Binet Scale were in the high 50's.

Anderson found a correlation of .45 between scores on the Gesell at two years and *IQ*'s at five years. He found a correlation of .52 between scores on a scale of tests made up of items from the Gesell, the Merrill-Palmer, and the 1916 Stanford-Binet at two years and Stanford-Binet *IQ*'s at five years. Cunningham obtained a correlation of .24 between two-year intelligence as measured by the Kuhlmann-Binet and intelligence six years later as measured by the Stanford-Binet.

Driscoll reported constancy correlations for 2- and 3-year intelligence as measured by the Kuhlmann-Binet and Merrill-Palmer tests which were similar to the ones found in the present study, but in some cases the time interval between Driscoll's tests was apparently as short as two years for the three-year-old group and three years for the two-year-old group The information reported in regard to the interval between tests was that the children were re-examined at the age of "five years or older "

De Forest's correlation of 71 between Merrill-Palmer *IQ*'s for 2- to 3½-year-old children and retest *IQ*'s was higher than the ones found in the present study, but the interval between tests in De Forest's study ranged from one month to 88 months The correlations found in the present study for 2- and 3-year-old children are approximated by the correlations between initial Stanford-Binet *IQ*'s at three years and retest Stanford-Binet *IQ*'s which were reported by Katz, but the time interval between tests in Katz's study was only two years as compared with an interval of approximately 10 years in the present study.

The highest constancy correlations which have been reported in the literature for these early ages are those of Hallowell Correlations reported by Hallowell between clinical appraisals of children at two and three years and later *IQ*'s were 80 and 84, which are considerably higher than those found in the present investigation for these ages In Hallowell's study the median age at the time of re-examination was six years The time interval between examinations was, apparently, no more than half the length of the time interval obtaining in the present study, consequently the findings cannot be compared directly The results suggest, however, that clinical appraisals based on tests coupled with other observations permit more accurate prognosis of future mental development than do objective measures of intelligence alone.

The comparisons made above indicate that *IQ*'s obtained from the Revised Stanford-Binet Scale at the ages of two and three years have better predictive value of future intelligence (as measured by the Stanford-Binet) than do other objective indices of two- and three-year intelligence for which results have been reported. The results from one study, however, suggest that clinical appraisals at these early ages have greater predictive value than have either Stanford-Binet *IQ*'s or other objective indices

The evidence regarding the predictive value of the Revised Stanford-Binet Scale as compared with other objective measures of intelligence is not as clear for the 4- and 5-year level. Comparisons suggest, however, that its predictive value for these age levels is as good as or better than that of other tests previously studied In no other instance have constancy correlations been reported in which the interval between tests for these specific age levels was as long as in the present study. The correlations between California Preschool scores at four years and Stanford-Binet *IQ*'s three years later which were reported by Honzik are lower than those found in the present study, but Honzik's correlations for five-year California scores with Stanford-Binet *IQ*'s two years later are higher than those reported here. The correlations between California scores at four and five years and Stanford-Binet *IQ*'s four and five years later which were reported by Bayley are also higher than those found in the present investigation. Since the intervals between tests in the studies of both Honzik and Bayley were no more than half that in the present study, however, a direct comparison between the correlations is not warranted De Forest's correlation between Merrill-Palmer *IQ*'s at four and five years and Stanford-Binet *IQ*'s one to 88 months later is about 20 points lower than those found in the present investigation.

A second method of studying *IQ* constancy in addition to the method of correlation is the consideration of the obtained *IQ* differences. Table 5 presents frequency and percentage distributions for differences between initial *IQ* (Form *L IQ*, Form *M IQ*, and composite *IQ*) and retest *IQ* for the two age levels studied The mean arithmetic differences are indi-

TABLE 5

FREQUENCY AND PERCENTAGE DISTRIBUTIONS FOR DIFFERENCES BETWEEN RETEST *IQ*'S AND INITIAL FORM *L*, FORM *M*, AND COMPOSITE *IQ*'S

IQ differences*	2- and 3-Year Group						4- and 5-Year Group					
	Form *L*		Form *M*		Composite		Form *L*		Form *M*		Composite	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
35 to 39							1	1	1	1	1	1
30 to 34	3	6					2	2	2	2	3	3
25 to 29	1	2	2	4	3	6	3	3	4	5	2	2
20 to 24	2	4	2	4	1	2	5	6	4	5	2	2
15 to 19	1	2	4	8	2	4	4	5	5	6	6	7
10 to 14	5	10	6	12	4	8	4	5	7	8	8	9
5 to 9	6	12	3	6	9	17	16	19	15	17	13	15
—4 to 4	14	27	12	23	12	23	30	35	22	26	30	35
—5 to —9	8	15	10	19	5	10	7	8	13	15	10	12
—10 to —14	2	4	3	6	7	13	9	10	7	8	8	9
—15 to —19	3	6	3	6	4	8	3	3	5	6	2	2
—20 to —24	4	8	3	6	1	2	1	1	1	1		
—25 to —29			3	6	2	4					1	1
—30 to —34			1	2	1	2	1	1				
—35 to —39	1	2			1	2						
—40 to —44	2	4										
Total	52	102	52	102	52	101	86	99	86	100	86	98
Mean Diff	13.1		12.4		12.1		10.3		10.9		9.9	

*Retest *IQ* minus initial *IQ*

cated in the last row of the table. The figures for the 2- and 3-year group show that regardless of whether the initial *IQ* was based on Form *L*, Form *M*, or a composite of the two forms, the *IQ*'s of about one-fourth of the children changed less than five points over the 10-year period, the *IQ*'s of about one-half changed less than 10 points, and the *IQ*'s of about two-thirds changed less than 15 points About one-third varied by 15 points or more over the 10-year period The percentage that increased by more than five points approximated the percentage that decreased by more than five points, although there was a slight tendency for a greater percentage to decrease—especially noticeable in Form *M* *IQ*'s. The mean change for Form *L* was 13 points; for Form *M* and composite *IQ*'s it was 12 points

The figures for the 4- and 5-year group show that regardless of whether the initial *IQ* was based on Form *L*, Form *M*, or a composite of the two forms, the *IQ*'s of about one-third of the children changed less than five points over the 10-year period, the *IQ*'s of about three-fifths changed less than 10 points, and the *IQ*'s of about three-fourths changed less than 15 points About one-fourth varied by 15 points or more. Approximately 50 per cent more subjects increased in *IQ* by five or more points than decreased by the same amount The discrepancy in per cent increasing and decreasing may have been due either to an unrepresentative sample of cases in the present study, or to inadequacies in the standardization which caused children at age 4 and 5 to "test low," or those at 14 and 15 to "test high" The mean *IQ* change for Form *L* and Form *M* was 11 points, whereas for composite *IQ*'s it was 10 points.

In Table 5 it will be noted that in spite of the close agreement between the correlations for the two age levels (see Table 4), the percentages within similar limits of *IQ* changes are higher for the older group Furthermore, the mean *IQ* changes are lower for the older group. It will also be noted that for both age levels Form *L* gives slightly higher percentages within given limits of *IQ* changes than does Form *M*. This is true in spite of the slightly higher correlation and lower mean *IQ* change for Form *M* for the 2- and 3-year group. Form *L*, however, tends to give more extreme changes than does Form *M*.

The mean arithmetic differences given in the last row of Table 5 may be compared with mean changes which have been reported by other investigators Since not all of the previous investigators reported mean changes in *IQ*, the comparison between present and previous results must be incomplete The comparisons which can be made, however, support the conclusions drawn from a comparison of the correlations in the present study with those found in previous studies, namely, that the predictive value of the Revised Stanford-Binet Scale is as good as or better than that of other tests previously studied.

E. Summary and Conclusions

1. One hundred and thirty-eight children who had been examined with both forms of the Revised Stanford-Binet Scale when they were two to six years old were re-examined with Form *L* of the Revised Stanford-Binet Scale 10 years later

2 Fifty-two of these children had been initially examined at the ages of 2, 2½, 3, and 3½ years and constituted the 2- and 3-year group in the present study. Eighty-six of the children had been initially examined at the ages of 4, 4½, 5, and 5½ years and constituted the 4- and 5-year group

3 The correlations between initial *IQ*'s and retest *IQ*'s for the 2- and 3-year group were: initial Form *L*, 58; initial Form *M*, 67, initial composite, 66. Correlations between initial and retest *IQ*'s for the 4- and 5-year group were: initial Form *L*, .67, initial Form *M*, .63, initial composite, .67 It was noted that for both age groups the correlation for one of the single forms reached or exceeded the correlation for the composite The differences between the correlations for the two forms was found to lack statistical significance.

4 It was found that the *IQ*'s from the first administered test at the time of the initial examination had as high predictive value of *IQ*'s 10 years later as did the *IQ*'s from a test administered a few days after the first test.

5. Comparison of the correlations obtained in the present study with those found in formerly reported studies suggest that the Revised Stanford-Binet Scale is as good as or better than any other objective index in predicting the future intellectual functioning of a preschool child

6 Regardless of whether the initial *IQ* was based on Form *L*, Form *M*, or a composite of the two, the *IQ*'s of about one-fourth of the 2- and 3-year group and one-third of the 4- and 5-year group changed less than five points after 10 years The *IQ*'s of one-half of the 2- and 3-year group and three-fifths of the 4- and 5-year group changed less than 10 points. The *IQ*'s of two-thirds of the 2- and 3-year group and three-fourths of the 4- and 5-year group changed less than 15 points

7 The mean *IQ* changes for the 2- and 3-year group were 12 points

for Form *M* and composite *IQ*'s and 13 points for Form *L IQ*'s, the mean *IQ* changes for the 4- and 5-year group were 10 points for composite *IQ*'s and 11 points for Form *L* and Form *M IQ*'s. These mean changes compare favorably with the mean changes, where reported, found in previous studies.

Although the number of cases in this study was not large, the conditions which obtained in the selection of subjects and the representative character of this sample group strengthen the value of the results. It is consequently justified to make certain preliminary generalizations regarding Revised Stanford-Binet *IQ*'s obtained at the preschool level. The results are encouraging, but at the same time point to the need for caution in the use of *IQ*'s at these early ages. An individual *IQ* obtained prior to the age of six years must be interpreted with discretion. The chances are about 1 in 4 if the child is four or five years old, and 1 in 3 if the child is two or three years old, that the *IQ* will change 15 points or more in the following 10 years. On the other hand, 3 out of 4 children aged four or five and 2 out of 3 children aged two or three will show a change of less than 15 points after 10 years. The implications depend on where the emphasis is put. Certainly for group prediction the preschool *IQ* has an important place. An individual preschool *IQ* which is unsupported by supplementary data, however, must not be used indiscriminately for making long-time predictions.

REFERENCES*

1. ANDERSON, J. E. The prediction of terminal intelligence from infant and preschool tests. *Yearb. Nat. Soc. Stud. Educ.*, 1940, **39** (I), 385-403.
2. ANDERSON, L. D. The predictive value of infancy tests in relation to intelligence at five years. *Child Devel.*, 1939, **10**, 203-212.
3. BAYLEY, N. Factors influencing the growth of intelligence in young children. *Yearb. Nat. Soc. Stud. Educ.*, 1940, **39** (II), 49-79.
4. ———. Mental and motor development from two to twelve years. *Rev. Educ. Res.*, 1939, **9**, 18-37.
5. ———. Mental growth during the first three years. *Genet. Psychol. Monog.*, 1933, **14**, 1-92.
6. ———. Mental growth in young children. *Yearb. Nat. Soc. Stud. Educ.*, 1940, **39** (II), 11-47.
7. ———. Studies in the Development of Young Children. Berkeley: Univ. California Press, 1940.
8. BAYLEY, N., & JONES, H. E. Environmental correlates of mental and motor

---

*See also the following publication, which appeared after the present article was written: Goodenough, F. L., & Maurer, K. M. The mental growth of children from two to fourteen years, a study of the predictive value of the Minnesota Preschool Scales. *Univ. Minn. Child Welf. Monog. Ser.*, 1942, No. 20.

development a cumulative study from infancy to six years *Child Devel*, 1937, **8**, 329-341.

9. Brown, R R. The time interval between test and re-test in its relation to the constancy of the intelligence quotient *J. Educ Psychol*, 1933, **24**, 81-96

10. Cunningham, B V. Infant *IQ* ratings evaluated after an interval of seven years *J Exp Educ.*, 1934, **3**, 84-87.

11. De Forest, R A study of the prognostic value of the Merrill-Palmer scale of mental tests and the Minnesota preschool scale *J Genet Psychol*, 1941, **59**, 219-223

12. Driscoll, G. P. The developmental status of the preschool child as a prognosis of future development *Child Devel. Monog.*, 1933, No 13

13 English, H. B, & Killian, C D. The constancy of the *IQ* at different age levels *J. Consult Psychol.*, 1939, **3**, 30-32

14 Foran, T. G The constancy of the intelligence quotient; a review *Cath Univ Amer Educ Res. Bull*, 1926, **1**, No. 10.

15 Garrett, H E Statistics in Psychology and Education New York Longmans, Green, 1932.

16 Gesell, A. The appraisal of mental growth careers. *J Consult. Psychol*, 1939, **3**, 73-75

17 ——— The Mental Growth of the Preschool Child. New York Macmillan, 1925.

18. ——— The stability of mental-growth careers *Yearb Nat. Soc. Stud. Educ*, 1940, **39** (II), 149-160.

19 Gesell, A, Castner, B M, Thompson, H., & Amatruda, C. S. Biographies of Child Development. New York. Hoeber, 1939

20. Goodenough, F L The relation of the intelligence of preschool children to the education of their parents. *Sch. & Soc.*, 1927, **26**, 54-56.

21. Goodenough, F L., & Anderson, J. E. Experimental Child Study New York Century, 1931.

22. Hallowell, D. K. Stability of mental test ratings for preschool children. *J Genet Psychol*, 1932, **40**, 406-421

23. ———, Validity of mental tests for young children. *J Genet. Psychol.*, 1941, **58**, 265-288.

24. Honzik, M. P. Age changes in the relationship between certain environmental variables and children's intelligence. *Yearb Nat Soc. Stud Educ.*, 1940, **39** (II), 185-205.

25 ——— The constancy of mental test performance during the preschool period *J Genet. Psychol*, 1938, **52**, 285-302.

26. Katz, E. The constancy of the Stanford-Binet *IQ* from three to five years. *J. of Psychol*, 1941, **12**, 159-181

27. Kawin, E Children of Preschool Age Chicago: Univ Chicago Press, 1934

28. Lowell, F E. A study of the variability of *IQ*'s on retest. *J. Appl. Psychol.*, 1941, **25**, 341-356

29. Macfarlane, J W. Studies in Child Guidance. I Methodology of data collection and organization. *Monog. Soc. Res. Child Devel*, 1938, **3**, No. 6

30. McNemar, Q The Revision of the Stanford-Binet Scale: An Analysis of the Standardization Data Boston Houghton Mifflin, 1942.

31. Meek, L H, & Jersild, A. T Mental development from two to twelve years *Rev Educ Res*, 1936, **6**, 17-48

32. Merrill, M A The significance of *IQ*'s on the Revised Stanford-Binet Scales. *J. Educ. Psychol*, 1938, **29**, 641-651

33 NELSON, V L., & RICHARDS, T. W Studies in mental development. I Performance on Gesell items at six months and its predictive value for performance on mental tests at two and three years. *J. Genet Psychol*, 1938, **52**, 303-325

34 ——— Studies in mental development: III Performance of twelve-months-old children on the Gesell schedule and its predictive value for mental status at two and three years *J Genet. Psychol*, 1939, **54**, 181-191

35 NEMZEK, C. L The constancy of the *IQ* *Psychol Bull*, 1933, **30**, 143-168

36. RICHARDS, T. W. Individual variations in *IQ* and analysis of concomitant factors *Psychol. Bull*, 1940, **37**, 442-443

37 STROUD, J B. Application of intelligence tests *Rev Educ Res*, 1941, **11**, 25-41

38. TERMAN, L M, *et al* The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence Baltimore Warwick & York, 1917.

39 TERMAN, L. M, & MERRILL M A Measuring Intelligence A Guide to the Administration of the New Revised Stanford-Binet Tests of Intelligence New York: Houghton Mifflin, 1937

40 THORNDIKE, R L "Constancy" of the *IQ*. *Psychol. Bull*, 1940, **37**, 167-186

41 ——— The effect of interval between test and retest on the constancy of the *IQ*. *J Educ Psychol*, 1933, **24**, 543-549.

42 TRELOAR, A. E Elements of Statistical Reasoning New York Wiley, 1939.

43 UPDEGRAFF, R The determination of a reliable intelligence quotient for the young child. *J Genet Psychol*, 1932, **41**, 152-166.

44. WALLIN, J. E. W. The results of multiple Binet retestings of the same subjects: The educational implications of variation of test performance *J. Genet Psychol*, 1940, **57**, 345-391

Department of Psychology
Stanford University
Stanford University, California

The Journal of Genetic Psychology, 1944, 65, 219-226

EARLY INDIVIDUAL DIFFERENCES IN VISUAL AND MOTOR BEHAVIOR PATTERNS A COMPARATIVE STUDY OF TWO NORMAL INFANTS BY THE METHOD OF CINEMANALYSIS*

The Clinic of Child Development, Yale University

Louise Bates Ames

A. Subjects

A photographic project at the Clinic of Child Development has recorded in detail the monthly behavior days of six normal infants throughout their first year of life (5) Each infant's behavior was continuously observed for one full day each month. Cinema records were made both indoors and outdoors, most of the records being taken within a domestic studio unit which could be converted into a kitchen (with bath), a playroom, bedroom, or living room as required. The normal behavior day of the infant constituted the central and governing objective for this photographic program The main phases of the day's cycle which were most consistently recorded included: sleep, feeding, bath and toilet, play, floor behavior, and social behavior. Normative examinations were also given and photographed at several of the monthly age levels

These children returned to the Clinic for re-study at the age of 5 and 12 years They were given detailed behavior examinations and were photographed during these and also during naturalistic and social situations Two of these children, whose demeanor was judged to be somewhat dissimilar on the basis of casual observation, were chosen as subjects for the present study of visual and motor behavior.

General Characterization

A detailed analysis of the individualities of these two boys, from 8 weeks through 5 years, has been worked out elsewhere (3). Consistent and marked differences between them were found to express themselves in the first few months of life, and continued to be observable through the fifth year A brief summary characterization of these two subjects quoted from this earlier study, follows:

---

*Accepted for publication by Arnold Gesell of the Editorial Board, and received in the Editorial Office on January 23, 1943

> As early as the ages of 8 and 12 weeks the highly dynamic personality of Boy *D* made a strong impression even when observed only through the medium of the cinema. After a full year of recorded behavior had been subjected to inspection the following adjectives were used to characterize his individuality quick, active, happy, friendly, well-adjusted, vigorous, forceful, alert, inquisitive . He was able to shift quickly in his emotional response from smiling to crying and from crying to smiling to achieve a desired end. At the age of five years, likewise, his emotional reactions are labile and versatile He is facile in changing his emotional responses He is highly perceptive of emotional expressions in others and, correspondingly, highly adaptive in social situations. With this emotional alertness he shows a relatively vigorous detachment from his mother as well as affection for her. He is not given to persisting moods.
>
> Boy *B* presents a different constellation of characteristics Although by no means emotionally shallow he is by comparison less vivid, less expressive, more self-contained than Boy *D* He is a sturdy, deliberate, moderately sociable, friendly child whose characteristics were evident at one year as well as at five years At five years he still shows motor ineptnesses and inhibitions comparable to those displayed in infancy. Boy *D*, on the other hand, has given consistent evidence of superior motor coordination from an early age in postural control, locomotion, and manual dexterity.

These descriptions were also to a considerable degree characteristic of the subjects at 12 years of age, at which time it seemed evident that Boy *B* was and had been consistently from infancy of a viscerotonic constitution and temperament whereas Boy *D* was definitely somatotonic (7).

As to *laterality*, both children are dominantly right handed at all ages. Boy *D* is also right footed. Boy *B* tends to use his left foot predominantly.

The rate of development for Boy *B* and Boy *D* was very similar during the whole 12 year period during which they were observed Boy *D* at each age rated slightly higher. Thus at 40 weeks Boy *B*'s maturity rating was 44 weeks, Boy *D's*, 44+ weeks; at 24 months Boy *B*'s rating was 30— months, Boy *D*'s 30+ months, at 60 months, Boy *B*'s rating was 72 months, Boy *D*'s 76 months Intelligence quotients at 12 years were 117 and 116 respectively.

B. PROCEDURE

Cinema records for each separate behavior situation (i.e, bath, feeding, floor play, lap play) at each monthly age level were analyzed, and a written description made of the outstanding activities and responses of the child to each situation. These descriptions were then reviewed and each was classified as to whether *regard only* (for some given stimulus object since the child would presumably be regarding *something* whenever his eyes were open), *regard plus contact,* or *contact only,* prevailed. Then for each child the percentage of situations at each age level when each type of behavior prevailed was obtained and graphed (Figure 1).

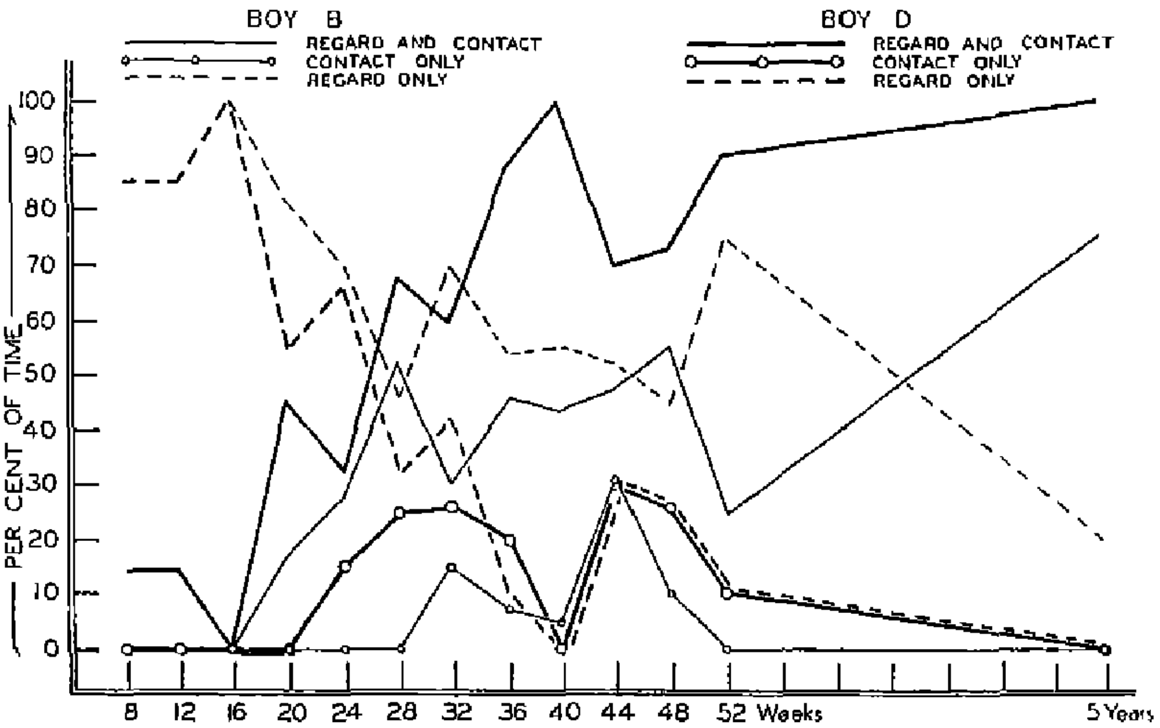


FIGURE 1

COMPARATIVE GRAPH SHOWING RELATIVE DISTRIBUTION OF REACTIONS OF BOY *B* AND BOY *D* TO NATURALISTIC BEHAVIOR SITUATIONS (SEE TEXT)

A more specific cinemanalysis measured the exact amount of time in seconds spent by each subject in regard for, contact with, and combined regard and contact with a stimulus object (a mirror) at 36 and 48 weeks.

And lastly the comparative speeds or tempos of movement for the two subjects were determined (by cinemanalysis) for the age range from 1 through 12 years, in both prehensory and locomotor situations.

C. FINDINGS

1. *Individual and Age Differences in Amounts of Regard and Contact*

Graphing of the percentage of all behavior situations (i.e., bath, feeding, play) at each age level when, for each subject, regard, contact, or regard plus contact prevails, shows clearly the marked and significant differences in the visual and motor patterns of these two subjects. Boy *B*, it will be observed, consistently spends more time in merely regarding a stimulus object than in contacting it At only three age levels does regard plus contact equal or surpass regard only. Contact only does not predominate at any age.

Boy *D* makes a very different response. From 8 through 24 weeks he spends the most time in merely regarding a stimulus object From then on, at *every* age level, regard plus contact markedly exceeds regard only. There is an unusually small amount of regard only. Contact only exceeds regard only at 36 and 52 weeks.

The age trends observed are of a somewhat similar nature for both children. Both *regard only* through the first 24 weeks and for both, at 28 weeks, contact plus regard for the first time surpasses regard only. From 28 weeks on, however, individual differences assert themselves, Boy *B* for the most part regarding more than he contacts, Boy *D* at all ages regarding with contact much more than he merely regards.

It will be of interest to compare these trends with the eye-hand responses of children from our normative research group (6) The responses of seven normative infants to the cube, pellet, and bell situations of the developmental examination have been determined (by cinemanalysis) for each monthly age level from 16 weeks to 60 weeks, and at yearly age levels thereafter up to 10 years of age The percentages of time spent in contact only, regard

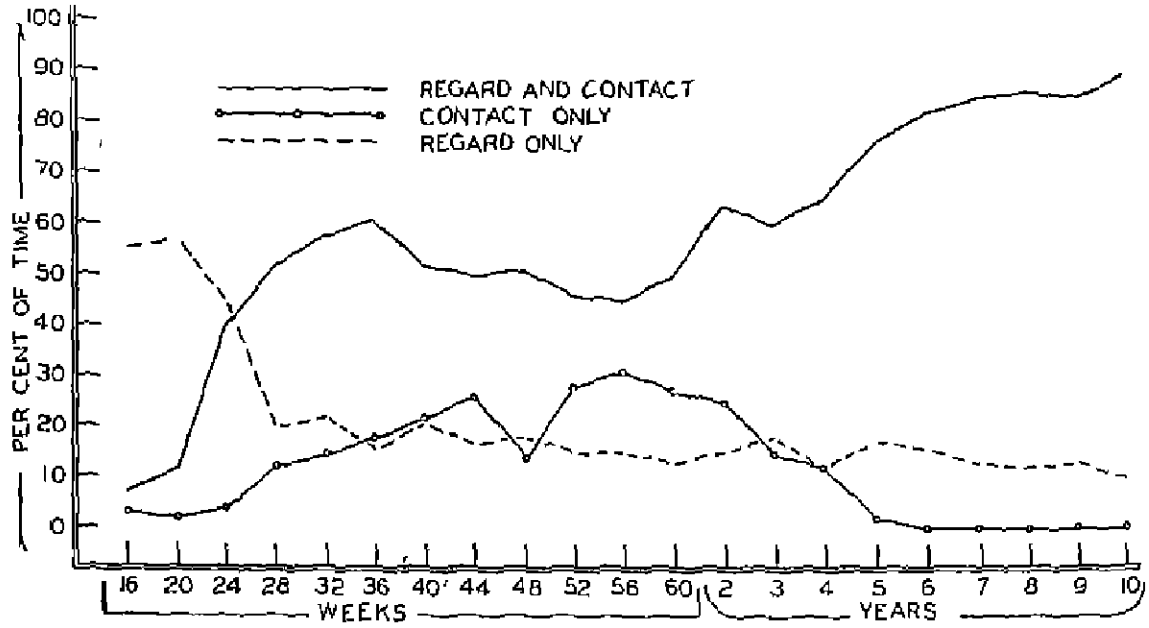


FIGURE 2
AVERAGE AGE TRENDS OF A GROUP OF SEVEN NORMATIVE INFANTS SHOWING DISTRIBUTION OF TIME DURING NORMATIVE BEHAVIOR SITUATIONS

only, and in regard plus contact have been determined, averaged for the seven cases and graphed (Figure 2).

This graph indicates that the normative trend of eye-hand patterns is as follows. Regard only predominates markedly at 16 and 20 weeks; predominates slightly at 24 weeks At 28 weeks *and thereafter* regard plus contact prevails conspicuously Regard only, and contact only occur less and show clear age trends Regard only usually prevails over contact only, though contact only exceeds regard only at 36 and 44 weeks, from 52 to 60 weeks, and at 2 years

Comparing the responses of normative cases with those of the present subjects, we find that *no normative child approximates the behavior of Boy B,* since in all normative cases regard and contact is the predominant behavior by 28 weeks and almost without exception thereafter. Boy *B* exhibits a definitely excessive amount of regard only; and much less than the usual amount of regard plus contact, than do normative cases.

Boy *D*'s response is much more like that of the normative group as a whole, and also of several individual normative cases For Boy *D,* regard only prevails through 24 weeks, and from then on, regard plus contact prevails. As with normative cases, between 36 and 60 weeks contact only exceeds regard only on several occasions; but unlike some individual normative cases, contact only *never* exceeds regard plus contact.

2. *Cinemanalysis of Responses to Mirror Situations*

An extremely objective treatment of some of the data on Boys *B* and *D*—actual counting of cinema frames—substantiates the conclusions already reached by descriptive analysis. Several mirror situations were chosen for cinemanalysis, at 36 and at 48 weeks In one situation at each age the child lay prone, about four feet from a mirror on the floor At 36 weeks he was also given an opportunity to lie, and later to sit upright, directly before the mirror and within reach ot it

The situations were equated as to exact relation of child to mirror, length of situation, absence of distracting stimuli, etc. At the earlier age, only Boy *D* had been observed to creep, at 48 weeks both children were expert creepers.

As Table 1 indicates, at both ages Boy *B*'s regard exceeded his contact; Boy *D*'s contact exceeded his regard Also, Boy *B*, when out of reach of the mirror lay quiet, regarded the mirror, and made no apparent effort to reach it He appeared to be undisturbed by the fact that it was out of reach.

TABLE 1
RESPONSES TO MIRROR SITUATION

Mirror out of reach		36 weeks behavior *Moves toward*	*Regards*		
First Presentation	Boy B	0%	76%		
	Boy D	100%	100%		
Second Presentation	Boy B	0%	40%		
	Boy D	100%	100%		
		Regards only	*Contacts*	*Regards + contacts*	*Does not regard*
Sits by Mirror	Boy *B*	57%	36%	36%	5%
	Boy *D*	0%	100%	93%	7%
Lies before Mirror	Boy *B*	74%	6%	6%	17%
	Boy *D* (1)	0%	100%	100%	0%
	Boy *D* (2)	0%	100%	100%	0%
		48 weeks behavior			
		Moves toward	*Regards*	*Contacts*	*Regards + contacts*
Mirror out of Reach	Boy *B*	0%	62%	0%	0%
	Boy *D*	62%	100%	0%	0%
Mirror within Reach	Boy *B*	(Does not reach mirror)			
	Boy *D*		73%	100%	73%

Boy *D*, on the other hand, spent most of the time when out of reach in moving toward the mirror

3. *Psycho-motor Tempo*

With regard to speed or tempo of movement Boy *D* is, for each measurable type of movement, definitely quicker than Boy *B*. Several hundred feet of cinema records were available for measuring the prehension time of these two boys in controlled developmental examination situations during the first year of life. The interval consumed between the zero moment of reaching to the moment of grasp of the test object was computed by counting cinema frames. Each frame has a time value of .05 seconds The average prehension time of Boy *B* at one year of age was .60 seconds; of Boy *D* only 40 seconds At 5 years of age, Boy *B* built a tower of 10 cubes in 23 seconds, Boy *D* in 16 seconds At 12 years, Boy *B* built a tower of 10 in 43 seconds, Boy *D* in 16 seconds

Creeping speed, like prehension time, was measured by counting frames, in this instance the number of frames which recorded a forward movement of a hand or of a leg. A creep advance represents a single cycle of progression accomplished by a single forward placement of each of the four members. The total time for a creep advance was determined. For Boy *B* this time was 1.60 seconds; for Boy *D* only .70 seconds, less than half as long. In prone progression we observe between the diagonal forward thrusts of an arm and the contralateral leg, a palpable pause which can be ascertained and measured by cinemanalysis. For Boy *B* the average time of this pause was 1.65 seconds, for Boy *D*, .65 seconds.

Furthermore, so far as the time for reaching each of the sequential stages of prone progression, Boy *B* was invariably older when he reached any given stage than was Boy *D*.

Since previous research (2) has suggested that infants and young children habitually maintain consistent speeds of movement relative to each other, we conclude that Boy *D* is not only more given to contacting any object in question than is Boy *B*, but that in his contacting he habitually moves more rapidly than does Boy *B*.

D. SUMMARY

1. The visual and motor patterns of two normal boys were studied in detail by means of cinema recordings of behavior taken at monthly intervals from 8 to 56 weeks and again at 5 and 12 years of age. Subjects were photographed in natural life situations (feeding, bath, play, social, etc.) as well as in the normative examination.

2. Behavior situations were classified as to whether regard, regard plus contact, or contact only (for a given stimulus object) prevailed.

3. Of these two subjects it appeared that in one *regard only* predominated at nearly every age level whereas in the other *regard plus contact* was the commonest type of response—*regard only* occurring very infrequently after 24 weeks.

4. Typical normative age trends for visual and motor patterns, presented graphically, indicate that, in typical normative infants, regard only predominates through 24 weeks and that at 28 weeks and thereafter regard plus contact prevails. Thus the relative amounts of regard and contact exhibited by any individual must depend to some extent on his age as well as on his individuality.

5. The two subjects were shown to be very similar in developmental rating from 40 weeks through 12 years, the child who exhibited the most contact rating slightly higher than the other. He was also found to move the more rapidly.

6. It appears that quantitatively measurable differences in the relative amounts of regard and of contact commonly characterize the behavior of certain infants The amounts of regard and of contact to some extent follow normative age trends, but vary in a characteristic manner for any individual child. Thus in some children regard behavior predominates over contact, and in others the opposite is true. Such differences are significant in understanding individuality. They should also be taken into account in rating developmental examination or other behavior, that we may not underestimate the abilities of a child who characteristically regards more than he contacts

REFERENCES

1. AMES, L. B. The sequential patterning of prone progression in the human infant *Genet. Psychol Monog.*, 1937, **19**, 409-460.
2 ——. The constancy of psycho-motor tempo in individual infants. *J. Genet. Psychol.*, 1940, **57**, 445-450.
3 GESELL, A. (WITH L B AMES). Early evidences of individuality in the human infant. *Sci Mo.*, 1937, **45**, 217-226
4 ——— & AMES, L B The ontogenetic organization of prone behavior in human infancy. *J Genet Psychol*, 1940, **56**, 247-263
5 ——— *et al* An Atlas of Infant Behavior. A systematic delineation of the forms and early growth of human behavior patterns (2 vols. 3,200 action photographs) New Haven Yale Univ Press, 1934. Pp. 922
6. ——— & THOMPSON, H The Psychology of Early Growth Including norms of infant behavior and a method of genetic analysis. New York: Macmillan, 1938 Pp. ix+290
7. SHELDON, W H (WITH STEVENS, S. S) The Varieties of Temperament. New York Harper, 1942 Pp x+520

The Clinic of Child Development
Yale University
14 Davenport Avenue
New Haven, Connecticut

The Journal of Genetic Psychology, 1944, 66, 227-237

GENERALIZATION OF THE CONCEPT OF MIDDLENESS*[1]

Institute for Research in Child Psychology, Hunter College

V. Graham, T A Jackson, L. Long, and L. Welch

Studies of the perceptual and conceptual development of children have included experiments dealing with size (2, 4, 6, 8) and form (3, 5, 7). In each of these experiments a single characteristic of the stimulus as, for example, roundness, triangularity, or middlesizedness had to be abstracted and the subjects were given the opportunity to learn the correct response in every new arrangement of the stimuli with which they were confronted. This procedure creates a certain difficulty in the interpretation of the results since it is possible that the child is learning to respond to a specific stimulus rather than to that generalized feature of the stimulus which the experimenter had in mind In the experiment of Hicks and Stewart (2), for example, middlesizedness was the concept under consideration. Children were tested on their ability to select the middlesized box, when any three of a series of six boxes were presented Since the same box was not necessarily the middlesized one on successive trials the correct choice was independent of any particular box. Before each choice, however, the child was told that the toy he was to find would be under the middlesized box Thus the concept was verbalized for him and he had only to learn to apply it to single specific situations.

One step beyond the concept of middlesize, middle weight, or any other aspect to which the term "middle" may be applied, is the general quality of middleness The present experiment was designed to investigate the extent to which children can form the general concept of middleness without being given verbal cues, and the procedure in the critical series was planned to reduce additional learning to a minimum. In general the procedure consisted of establishing an initial response to *middlesizedness* and then observing the ability of the subjects to select stimuli of middle *density* The "size" stimuli were squares varying in area The "density" stimuli were circles of identical area, but containing a varying number of small white dots so that the ratio of the amount of black to the amount of white within circles of a constant area was varied.

---

*Received in the Editorial Office on February 1, 1943

[1] From the Institute for Research in Child Psychology, Hunter College The authors are indebted to Robert Chin for his assistance in standardizing the procedure used in this research project

In addition to being interested in determining whether the initial response to middlesizedness can be set up and the extent to which the concept of middleness can be generalized, several other related problems were also considered. (*a*) the relationship between age and the establishment of the initial response, (*b*) the relationship between age and the ability to form a general concept of middleness, and (*c*) the relationship between the first choice made by the child on the post-training series and subsequent responses in that series

APPARATUS

The apparatus consisted of a large wooden box, similar in appearance to a toy piano and the subject was seated before it as if it were a piano In the position corresponding to the keyboard were three square openings with doors formed by hinged white bakelite plates which could be opened by a downward pressure. By means of a switch the experimenter could prevent any one, or all three doors, from opening. A shelf beneath these doors held three dishes, one immediately below each door. When the door opened the subject could reach into the dish and obtain a small piece of candy placed there A mirror system was devised which made it possible for the experimented to project the stimuli upon the doors from a projector placed behind the apparatus Each projector slide contained three geometrical figures and the apparatus was adjusted until one of these figures fell on each of the three doors

In a preliminary experiment the "just noticeable difference" was obtained for two series of stimuli, which varied in area and density. A series of nine squares of black paper (mounted on white backgrounds) varying in size by a constant amount were presented to college students The middle stimulus was compared with each of the other eight stimuli in haphazard order as is the custom in the method of constant stimuli (1). As each pair of stimuli was presented the subjects indicated whether the variable was "larger than" or "smaller than" the standard. The difference limen was then calculated and was expressed as a percentage of the standard. A value equal to 10 times this percentage was selected as the extent to which the middle stimulus of the middlesize experiment should differ from the smallest stimulus. The proportional difference between the largest stimulus and the middle stimulus was the same as that between the middle stimulus and the smallest stimulus. Another preliminary series of nine stimuli varying in density but constant in area was presented to college students The stimulus of middle density was used as the standard and each variable was paired twice with the standard The subjects indicated each time whether the variable was "more dense" or "less dense" than the standard As in the procedure outlined above for the selection of the size stimuli, the difference limen was calculated and expressed as a percentage of the standard This value, multiplied by 10, was the extent by which the stimulus of middle density differed from the stimulus of least density. The stimulus of greatest density was selected so that the proportional difference between it and the stimulus of middle density would be the same as that between the stimulus of middle density and the stimulus of least density This method of equating difference limens for area and density series is of course accepted psychophysical procedure, but a measure of the "goodness" of the equating is not available

Twelve slides, each of which contained three squares with the proportional variation in area described above, were made The slides were exact duplicates except that the position occupied by the smallest, middle, and largest squares varied in a chance manner. These slides were used in the training period and will be referred to hereafter as the stimuli of the training series Six other slides, each of which contained three circles with the proportional variation in density described above, were made. They were exact duplicates except for the position of the circle of least, most, and middle density This set of slides will be referred to as the stimuli of the density series Another set of six slides which combined the area and density stimuli was used in a third series which will be called the area-density series, i e., the largest square of the area series was given a density equal to that of the stimulus with the greatest density in the density series, and similarly for the middle size square and the stimulus of middle density, and for the smallest square and the stimulus of least density. The 12 stimuli from the density and area-density series were presented to the subject in random order and as such formed the stimuli of the critical or test series

PROCEDURE

After the child was seated before the apparatus the following instructions were given.

> I am going to show you some pictures. You will see them, three at a time, one on each of these doors (pointing to the doors). Sometimes the doors open and sometimes they do not (demonstrating) You want to

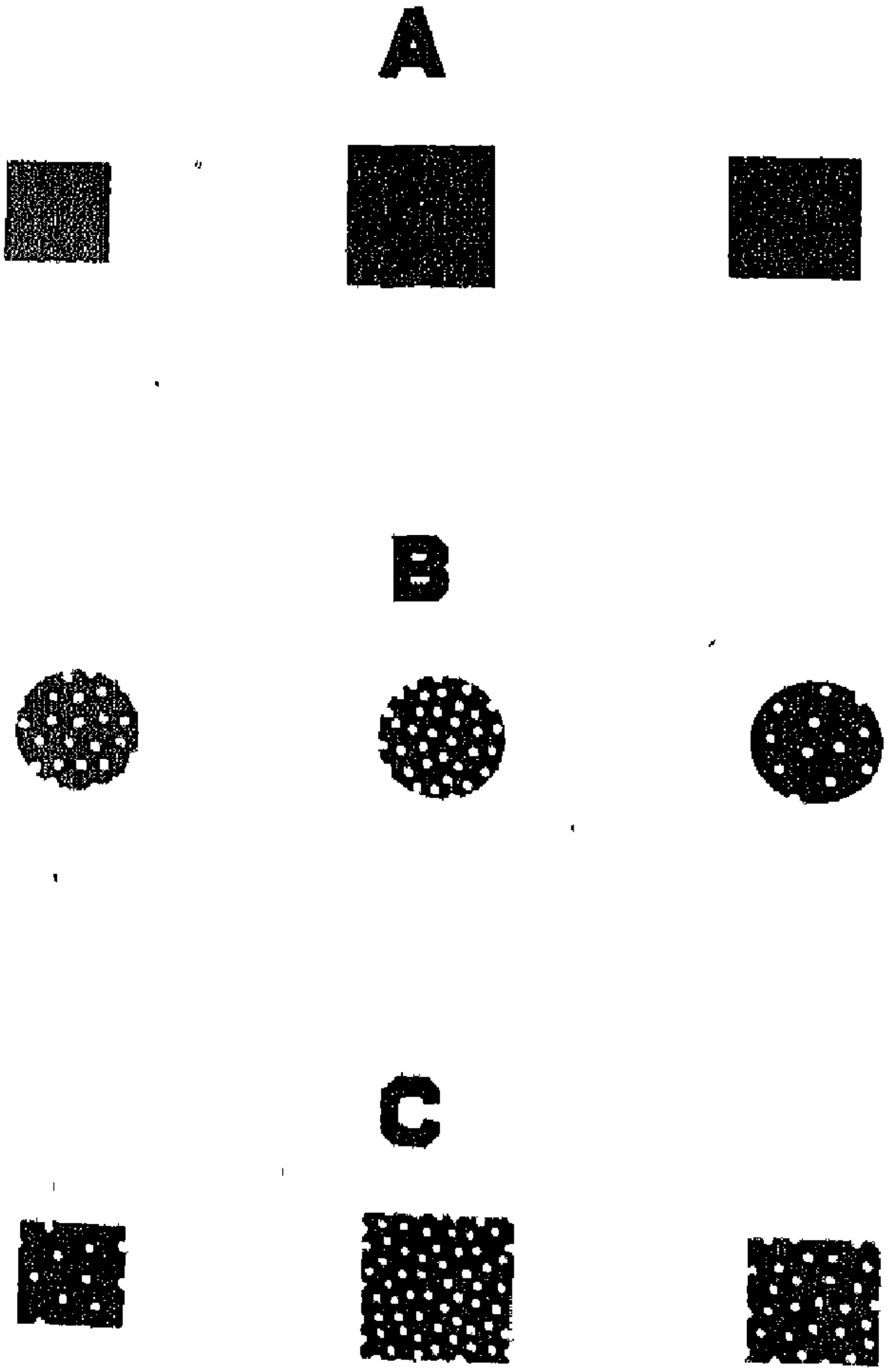


FIGURE 1

Typical Stimuli of Each of the Three Series

A—(training series, *B*—density series (critical series); and *C*—area-density series (critical series).

> find out what will make one of the doors open, because if you can open it, you may have the candy in the dish under the door. (Show candy to child) Now remember you will see one picture on each door *One* of these pictures means that the door on which you see it will open, and you should try to find out which picture it is Each time that I show you three pictures, see if you can find out which one will open the door. If you are right, reach down and get the candy Then we will try three more pictures If you are wrong you will have three more pictures and another chance to get the candy You may touch only one door each time Do you understand what to do?

If the child asked any questions, he was given no further information or hints, but an effort was made to repeat word for word whatever part of the instructions he had not understood. As soon as the experimenter felt that the child understood the task, the stimuli of the area series were presented until the subject fulfilled the criterion of 12 consecutive correct responses or until 100 responses were made to these stimuli Throughout this training period the apparatus was adjusted so that the door under only the middlesized square would open When 12 consecutive correct responses had been made it was assumed that the concept of middlesizedness had been established. If this criterion was not fulfilled within 100 trials no further training was given nor were the test stimuli presented

If the criterion of the training series was fulfilled the stimuli of the test series were presented All doors were unlocked and candy was placed under each door throughout the test series Naturally the child was not informed of this change in procedure The following additional instructions were given·

> Now you will be shown some more pictures and again you are to pick which door is going to open Try to figure out which picture will open the door You may touch only one door each time.

After the instructions the child was presented with three stimuli from the density series and three from the area-density series in random order. If he selected the middle stimulus on all six occasions, he was immediately presented with the remaining group of six stimuli, three from each of the two test series. If he again selected the middle stimulus on all occasions the experiment was terminated If, on the other hand, he failed to select a middle stimulus when the test series was presented, he was returned to the training series. The stimuli of the latter series were presented until the criterion of 12 consecutive correct responses was fulfilled or until the cumulative total of training trials equalled 100. If the criterion was fulfilled he was given the second set of six test stimuli. If he made six responses to the middle stimulus he was then given the first set of six test series (viz., the ones he had been tested on immediately after the initial training). If he selected the middle stimulus each time, the experiment was ended, since he had made 12 consecutive responses to the middle stimulus To summarize, if mistakes were made on the test series the subject was retrained until the criterion of the training series was again fulfilled; he was then retested; if he failed to select the middle stimulus in any of the 12 test trials, he was retrained, and then retested. This procedure was continued, if necessary, until the subject had been retrained three times—that is, he was given four opportunities to generalize If he did not the experiment was terminated.

SUBJECTS

Results were obtained from 75 children of the Hunter Elementary School. The age range of the subjects was from 84 to 117 months. The children were divided into VII, VIII, and IX year groups with 25 children in each group. Each of the subjects was tested individually and at one experimental session of not more than 60 minutes Motivation was excellent

DISCUSSION OF RESULTS

The results of the experiment suggest that the children can be divided into three groups on the basis of their performance The first group (referred to as the failure group) includes those children who were considered to have failed to learn the initial discrimination; i e, they either failed to fulfill the criterion of 12 consecutive correct responses within 100 trials or failed to develop a consistent response in the critical series after fulfilling the criterion in the training series The second group (referred to as the partial success group) includes those subjects who passed the training series and were able to formulate a definite system of responding in the critical series although that system was not a generalization of middleness. The third group (referred to as the complete success group) includes those children who systematically selected the middle stimulus in the critical series Each of these groups has been further divided and the sub-groups are described in Table 1 This table also presents the number of children falling into each category according to age. This material is summarized in Table 2 where only the three main groups are listed. It will be seen that a very marked relationship between age and the ability to generalize the concept of middlesizedness is indicated; for example, no nine-year-old child is included in the failure

TABLE 1

SUMMARY OF PERFORMANCE ON TRAINING AND CRITICAL SERIES

Group	Type of stimulus selected Training series	Critical series	Number of subjects in each group VII	VIII	IX
Failure					
a)	Inconsistent	(not given to subjects)	13	6	0
b)	Middle	Inconsistent	3	3	0
Partial success					
a)	Middle	Greatest	0	2	2
b)	Middle	Least	4	4	7
c)	Middle	Not the same for both series*	3	4	5
Complete success					
a)	Middle	Middle (no retraining)	1	2	3
b)	Middle	Middle (retrained)	1	4	8

*Type of stimulus selected by these children in

Area-density series	Density series	VII	VIII	IX
Middle	Least	0	2	2
Middle	Greatest	0	0	1
Middle	Uncertain	3	2	1
Greatest	Least	0	0	1

TABLE 2

RELATIONSHIP BETWEEN PERFORMANCE AND AGE

Performance in critical series	Number of subjects in each age group VII	VIII	IX	Total
Failure	16	9	0	25
Partial success	7	10	14	31
Complete success	2	6	11	19
Total	25	25	25	75
Chi square 24.25		$P<0.01$		

group, while 16 of the 25 seven-year-old children are included in this group. The relationship between age and success was found to be statistically significant when the chi square test was applied to the frequencies listed in Table 2 ($P<0.01$).

The records of the 19 children who did not fulfill the criterion of passing the training series were analyzed in an effort to determine the basis of selection. Patterns or sequences of responses based on position were occasionally observed. An example of such a sequence would be selecting the left, middle, and right doors on successive trials. A slightly different positional factor operated in the case of four other subjects in so far as each one responded much more frequently to one particular door than would have been expected on the basis of chance alone. Repetitive types of responses were also observed. No attempt has been made to quantify the extent to which these systems or hypotheses operated since the number of subjects failing was only 19.

An analysis was also made of the number of trials necessary for the subjects to fulfill the criterion of 12 consecutive correct responses in the training series. In this analysis only those subjects who fulfilled the criterion were included In Table 3 the average number of trials needed by the children in the three age groups are presented. The younger children required a greater number of trials to learn the initial discrimination than did the older ones, but the differences between the averages for the several age groups were found to be not reliable (Student's t-Test gave P values greater than 0.05).

TABLE 3

RELATIONSHIP BETWEEN AGE AND NUMBER OF TRIALS BEFORE FULFILLING CRITERION OF TRAINING SERIES

(Data from subjects *not* fulfilling criterion are omitted.)

	Age groups VII	 VIII	 IX
Number of subjects	12	19	25
Av number of trials	26 67	23 16	16 04

Previous experimental work in the animal as well as in the human field has suggested that the first response made in the critical series is of great importance in determining whether or not the subject has generalized from the training series. The close agreement between selecting the middle stimulus on the initial trial and on subsequent trials presents added evidence that these two measures can be used to supplement each other Fourteen of the 19 children in the complete success group selected the middle stimulus on the first trial, whereas only one of the 31 children who were classified as partially successful in generalizing selected the middle stimulus on the first trial (Table 4) The association between the first response and succeeding responses was found to be statistically reliable ($P<0\ 01$) and it can be concluded that the child who selects the middle stimulus on the first critical trial will ordinarily continue to select this stimulus throughout the experiment.

The development of systems classified as partially successful in generalizing is no doubt related to the influence that the initial response has on subsequent responses. For example, if the subject selects the figure of least

TABLE 4

RELATIONSHIP BETWEEN FIRST RESPONSE AND SUBSEQUENT RESPONSES IN CRITICAL SERIES

Performance on critical series	Number of subjects selecting various doors* on first trial Right	Middle	Left	Total
Partial success	24	6	1	31
Complete success	2	3	14	19
Total	26	9	15	50

Chi square 29.71 $P<0.01$

*Right door—smallest stimulus
Middle door—largest stimulus
Left door—middle stimulus

density when he is presented with figures differing in density he may continue to select this particular type of stimulus throughout the rest of the experiment, since all responses of the initial series are rewarded

A tendency was noticed for the younger children to select the stimulus on their right when the critical series was begun. This tendency is not evident in the data obtained from the children in the oldest group since the three doors were selected approximately an equal number of times on the initial trial of the critical series The frequencies presented in Table 5 deviate from a chance distribution ($P<0.05$) and it will be seen that the responses of the children in the seven- and eight-year groups are primarily responsible for the non-chance distribution.

TABLE 5

RELATIONSHIP BETWEEN AGE AND FIRST RESPONSE IN CRITICAL SERIES WITHOUT REGARD TO SUBSEQUENT RESPONSES IN CRITICAL SERIES

Age group	Number of subjects selecting various doors* on first trial of critical series Right	Middle	Left	Total
VII	8	3	1	12
VIII	11	3	5	19
IX	10	6	9	25
Total	29	12	15	56

Chi square 13.04 $P<0.05$

*Right door—smallest stimulus
Middle door—largest stimulus
Left door—middle stimulus

An inspection of the verbal comments made by the subjects suggests that the concept of middleness was a part of their vocabulary. For example, only five of the children who learned the initial discrimination failed to use terms such as "middlesized square," "medium one," or "middle one" These five referred to the middle stimulus as the "next to the biggest," "between the big and the little," or "next to the smallest" The explanation of the choice in the critical series showed more variety. Many of the children seemed to view the density and the density-area series as two separate problems; for example, one child said: "In the squares it is the middlesized one, in the circles it is the one with the middle number of dots." In a few cases the child's choices did not coincide with his verbalizations. An illustration of this can be found in the case of the child who said "in the square (series) it is middlesized, in the circle (series) it is the one in the middle," but yet he consistently picked the circle with the smallest number of dots. The clearest explanation was given spontaneously by an eight year old girl when first presented with the critical series· "these are different ones but it's the same trick; the medium one opens the door, it's the same in the plain and in the dotted"

Summary and Conclusions

A discrimination problem was presented to 75 children ranging in age from seven to nine years. In a training series the children were presented with three squares varying in area and were rewarded only when the square of middle area was selected. If the middle stimulus was selected on 12 consecutive trials a critical series was then presented and all responses were rewarded The critical series consisted of circles varying in density and squares varying in both area and density. The results indicate that the ability to form a generalized concept of middleness increases between the ages of seven and nine years A definite relationship between the first choice in the critical series and later performance in this series was found.

References

1 Guilford, J. P. Psychometric Methods. New York, McGraw-Hill, 1936 Pp 566

2 Hicks, J A, & Stewart, F D The learning of abstract concepts of size *Child Devel*, 1930, **1**, 195-203.

3 Long, L Conceptual relationships in children: the concept of roundness. *J. Genet. Psychol.*, 1940, **57**, 289-315

4 ———. Size discrimination in children. *Child Devel.*, 1941, **12**, 247-254.

5 Munn, N. L, & Stiening, B. R The relative efficacy of form and background in a child's discrimination of visual patterns. *J. Genet. Psychol*, 1931, **39**, 73-90.

6. Thrumm, M E The development of concepts of magnitude. *Child Devel.*, 1935, **6**, 120-140.

7. WELCH, L. The development of discrimination of form and area *J of Psychol.*, 1939, **7**, 37-54

8. ———. The development of size discrimination between the ages of 12 and 40 months *J. Genet Psychol*, 1939, **55**, 243-268.

Institute for Research in Child Psychology
Hunter College
New York City

The Journal of Genetic Psychology, 1944, **65**, 239-263

BINET *IQ* CHANGES OF ORPHANAGE PRESCHOOL CHILDREN. A RE-ANALYSIS*

Iowa Child Welfare Research Station, State University of Iowa

BETH L. WELLMAN AND EDNA LEE PEGRAM[1]

The purpose of this study is to present some new analyses of the data collected in an orphanage preschool project on Binet *IQ* changes of children subjected to orphanage and orphanage-preschool environments (2)

A SUMMARY OF THE ORIGINAL STUDY

The purpose of the original research study (2) was to determine the effects of preschool education introduced into the lives of a group of underprivileged children of average and below average intelligence. This was accomplished by setting up a preschool in which half of the children of preschool age in residence in the orphanage were enrolled The progress of the preschool children was compared with that of children of equal initial status who experienced the same environment in other respects but who did not attend preschool

Before the introduction of the preschool, all children of preschool age in the orphanage had been living in an environment in which routines were the main events of the day, with the time between filled mainly with sitting simply waiting

The children's school and cottage schedule after introduction of the preschool may be described in general The preschool started at eight o'clock. Vigorous play was usually planned at this time, followed by quieter activities, toilet, and finally tomato juice and cod-liver oil at about 9·30 o'clock Then came a rest period, followed by activities of a constructive kind, musical experiences, stories, or an excursion The children washed themselves at school in preparation for the noon meal and left for their cottages to go immediately to the main building at noon. After naps in the cottages, the children returned to the preschool where they had experiences with books and music and engaged in motor activity and constructive play At five o'clock the children were called for and then supper came soon after, with an early bedtime

---

*Received in the Editorial Office on February 2, 1943

[1] The writers are indebted to Dr Leon Festinger, Dr. Virginia B Knott, and Dr. Greydon Worbois for their helpful suggestions and services in computation.

The control children followed the same schedule of arising, going to bed, and going to meals. Their activities between times were handicapped by lack of equipment, space, and adult supervision. The duties of adults in charge were too heavy to allow for much encouragement of intellectual interest and initiative in the children Sitting still and being quiet were to them (and understandably so) more desirable traits than the normal running about and asking of questions of young children. Play materials were at a minimum.[2]

At the beginning of the preschool project, all children of preschool age in residence in the orphanage were divided into two groups matched on chronological age, mental age, intelligence quotient, sex, nutritional status, and length of residence in the orphanage prior to the opening of the project. Wherever possible the children were paired, one member to attend preschool and the other to be in the control group. In October, 1934, when the preschool opened, there were 21 pairs of children ranging in age from 18 months to five and one-half years. During the three years of the project, other children entered the orphanage and were included in one of the two groups, while others left for adoption or entered kindergarten During the course of the study an attempt was made to keep the groups as nearly equated as possible on the same bases as the original matchings. At the conclusion of the project, there had been 59 preschool and 53 control children included.

Analyses of effects of preschool attendance upon intelligence in the original study (2) were made on 46 preschool[3] and 44 control children. The other children could not be used because they did not have two intelligence tests, the reason being short residence in the project. Analyses were made in terms of cases; three groupings were made according to residence between tests, and a child was included as a case as many times as he fulfilled the residence requirement for a particular grouping, i.e, 1 to 199 days, 200 to 399 days, and 400 or more days between tests

The children were examined on either the 1922 Kuhlmann or the 1916 Stanford revision of the Binet scale. The number of examinations per child varied up to six tests, the number of tests being dependent mainly upon the length of time the child remained in the project. The examiners were staff members or research assistants in the Iowa Child Welfare Research Station. At the time of testing, except on initial tests, they knew which children belonged to the preschool and which to the control group, but did not know what *IQ* a given child had previously received There was no attempt to have a child assigned to the same examiner for retest

---

[2] A fuller description of the life of the control group will be found on pages 10-13 of the original study (2) and of the preschool group on pages 20-36

[3] In the process of re-analysis, it was discovered that one preschool child (child No. 35 in Table 2) who had two intelligence tests had been inadvertently omitted from the original study because of admission to the project late in its third year, he was included in the re-analysis, making the number of preschool children 47.

Analyses were made of changes in *IQ* according to initial *IQ* by 10-point classifications and length of residence, of changes of children above and below 80 *IQ*, of changes according to number of days of preschool attendance, and of changes in relation to percentage of preschool attendance.

Distinct differences in pattern of growth were found between preschool and control children For the group as a whole, there was no difference in trend between preschool and control children during the shortest residence period of approximately four months Subsequently, the preschool and control groups began to diverge and the divergence became accentuated as time went on Over the longest residence period of approximately 20 months, the preschool children (40 cases) gained 4 6 points in *IQ*, while the control children (65 cases) lost 4.6 points.

The difference in pattern of changes was more marked when the groups were subdivided according to initial *IQ* into 50 to 79 initial *IQ*, and 80 or above initial *IQ* Control children testing 50 to 79 *IQ* made about a constant 3-point gain in *IQ* over the three residence periods. Preschool children testing initially 50 to 79 made the same gain as control children over the shortest residence period, but made a gain of 7.7 points over the longest residence period. Preschool children testing 80 and above initially remained about constant while the control children lost markedly (6.6 points for 41 cases over nine months, 16 2 points for 26 cases over 20 months)

Within the preschool group changes were related to number of days of preschool attendance Each one hundred days' attendance resulted in more gain than the previous group. The changes were. 1 to 99 days, —1.2 points; 100 to 199 days, 3.6 points, 200 to 299 days, 5.1 points; and 300 or more days, 6.0 points

The conclusions made from the study were that this preschool made an effective contribution toward a better intellectual development of the children who attended, and that the contribution of the preschool appeared to be that of counteracting losses toward which the non-preschool environment of these children was constantly pulling

The findings of this study were not dissimilar to those of the only other study of environmental stimulation of orphanage preschool children of which the authors are aware, that of Barrett and Koch (1). In their study the children who were to attend preschool were from the Chicago Orphan Asylum, while the control children were drawn from this and three other orphanages which were similar in general equipment, physical care offered the children, daily routine, type and training of those in authority, and number of children handled by one caretaker. The average gains in *IQ* (Merrill Palmer scale) for the nursery school and control groups were, respectively, 20.6 points and 5.1 points. The authors concluded that nursery school activities and methods in some way raised the general accomplishment of the children

B. Plan of the Re-Analysis

This re-analysis will (*a*) present the basic data in convenient form for reference, (*b*) utilize for each individual only the initial and terminal Binet tests, and (*c*) make some new types of analysis.

The report to follow is divided into two sections. The first section presents analyses of variance between preschool and control groups according to preschool attendance and residence The second section presents correlations of *IQ* with residence and with preschool attendance

C. Composition of the Groups

Table *A* shows the equivalence in age and *IQ* of the 21 pairs of children who were in residence when the preschool was opened.

TABLE *A*

	Mean		Standard Deviation	
	Preschool	Control	Preschool	Control
Chronological age, months	44.6	44.4	16.6	16.8
IQ	82.3	81.2	12.3	11.6

In addition, nutritional status of the children in residence when the project was started was determined by two ratings, one based upon the medical examination and the other on anthropometric measurements. In both series the ratings were done on a five-point scale The mean ratings for the 18 pairs for whom these measures were available were 2.5 and 2.7 respectively for preschool and control groups on the physical measurements, and 3.1 and 2.8 respectively on the medical score.

The sexes were evenly distributed; each of the original groups included 12 boys and 9 girls. At the opening of preschool, the preschool children had been in residence in the orphanage 21.8 months and the control group 18.4 months

After the project had begun, new admissions during the three-year period were added to each of the groups with an attempt continually to equalize the two groups in relation to the factors considered in the original matching.

As was stated previously, although 59 children were admitted to the preschool group and 53 children were admitted to the control group at some time during the three years of the project, 12 of the preschool children and 9 of the control children did not remain long enough to receive two intelligence tests. They were of necessity excluded from a consideration of *IQ* changes. This left 47 preschool children and 44 control children who received two intelligence tests and who constituted the groups used in the study[4]

Some of the children were in the project for only a short time. There were three reasons for short residence in the project (*a*) placement for adoption, (*b*) age, and (*c*) late entrance to project. These factors appeared to operate about equally in control and preschool groups. Sometimes children whose *IQ*'s were relatively high (usually 90 or above) were removed for placement in adoptive homes The number of children removed from the control group for this purpose was 15, in addition to the 9 children who did not receive two intelligence tests, or a total of 24 The number of children removed from the preschool group for adoptive purposes was 12, in addition to the 12 who did not receive two intelligence tests, or a total of 24

Table 1 presents for each individual in the control group the number of days of residence from the date of his entrance to the project to the date of his last test, and his chronological age and *IQ* on each test[5] For a control child in the original matched groups residence during the project was considered as beginning on the date the preschool opened.

Table 2 presents for each individual in the preschool group the number of days of residence during the project, the number of days attendance in preschool, the percentage of preschool attendance, and his age and *IQ* on each test Residence during the project for a preschool child was considered as beginning on the date of his entrance to preschool. Percentage of preschool attendance was computed by dividing total days attendance in

---

[4] See previous footnote

[5] Only the initial and terminal tests have been used in the re-analysis The intervening tests are presented here for reference, since they were utilized in the original study (2).

TABLE 1

CONTROL GROUP

(Ages and IQ's on Successive Tests and Residence during Project.)

Child	Residence, days	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*
1	972	23	80	29	80	33	74	36	74	45	67	57	70
2	971	18	98	25	93	28	83	31	80	40	61	52	61
3	971	18	103	27	72	38	63	50	60				
4	970	27	71	33	66	39	66	48	71	55	79	61	79
5	963	24	75	28	78	32	88	34	84	43	84	56	75
6	859	22	83	26	83	29	75	38	63	50	60		
7	858	19	97	23	95	26	89	35	86	47	94		
8	839	25	73	31	63	35	58	38	65	47	64	55	62
9	686	55	60	58	79	66	67	73	68				
10	648	25	62	30	53	38	64	46	53				
11	645	45	57	53	65	61	59						
12	598	43	77	49	74	53	83	56	82	65	86		
13	596	32	74	37	74	41	70	43	73	52	88		
14	596	40	72	45	70	49	67	52	69	61	69		
15	594	37	61	44	73	48	83	51	82	60	83		
16	483	42	96	47	87	49	90	58	86				
17	470	42	76	51	75	56	82						
18	469	37	114	40	106	43	102	52	100				
19	466	49	93	52	77	55	76	64	73				
20	458	31	88	42	72	47	83						
21	442	31	85	39	70	45	71						
22	345	30	73	38	75								
23	335	17	103	29	85								
24	320	69	70	75	67	79	80	82	78				
25	320	59	84	64	91	69	87	71	85				
26	320	63	76	69	70	73	74	75	83				
27	320	55	91	58	97	62	100	65	95				
28	319	64	81	69	81	74	84	76	79				
29	266	54	67	62	84								

TABLE 1 (*continued*)

Child	Residence, days	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*
30	240	62	65	68	71	72	76						
31	217	61	89	65	94	69	87						
32	213	32	85	38	84								
33	192	44	91	48	108								
34	179	23	96	36	98								
35	174	60	70	61	85	64	97						
36	171	66	79	72	81								
37	118	57	95	62	94								
38	118	18	102	22	100								
39	114	18	100	23	92								
40	113	40	106	45	93								
41	113	37	86	42	103								
42	112	51	106	56	96								
43	111	61	85	67	101								
44	110	32	79	35	75								

TABLE 2

PRESCHOOL GROUP

(Ages and *IQ*'s on Successive Tests, Residence during Project, Preschool Attendance, and Percentage of Attendance.)

Child	Residence, days	Attendance	Per cent attendance	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*
1	970	572	59	31	66	36	72	40	67	43	63	52	77	64	75
2	970	543	56	25	76	31	75	35	70	38	73	47	68	59	81
3	840	474	56	24	98	27	84	29	89	39	97	51	94		
4	620	351	57	43	73	52	73	64	75						
5	596	296	50	37	68	43	67	47	85	50	92	59	85		
6	593	347	59	35	66	42	73	45	70	48	78	57	91		
7	588	292	50	21	94	26	85	28	72	31	74	41	89		
8	588	249	42	26	80	31	73	35	74	38	68	47	83		
9	587	314	53	38	99	42	91	47	94	50	96	59	105		
10	566	311	55	29	108	35	95	39	112	42	100	51	94		
11	544	247	45	22	92	29	78	41	94						
12	537	303	56	42	71	48	66	60	83						
13	479	218	46	29	85	33	82	37	82	39	81	45	89		
14	478	236	49	48	79	54	81	58	86	61	89	67	76		
15	478	267	56	46	104	52	119	55	120	59	122	64	109		
16	477	292	61	52	78	55	80	57	84	66	97				
17	468	166	35	22	84	26	75	38	71						
18	466	101	22	27	80	31	62	43	76						
19	422	205	49	53	94	56	86	60	85	65	92	69	90		
20	418	255	61	43	73	57	84								
21	432	188	44	25	74	37	59								
22	321	185	58	59	83	65	89	69	90	72	89				
23	319	187	59	65	75	72	78	76	84	79	76				
24	318	187	59	62	84	66	86	70	91	72	89				
25	316	179	57	29	80	34	93	38	84	41	87				
26	310	197	64	68	71	74	76	78	82	80	75				
27	269	110	41	27	85	33	80	36	80						
28	266	142	53	50	68	61	73								
29	264	148	56	55	87	61	98	64	94						

TABLE 2 (*continued*)

Child	Residence, days	Attendance	Per cent attendance	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*	Chronological age, months	*IQ*
30	261	165	63	52	77	61	89								
31	217	133	61	34	84	40	80	43	93						
32	189	102	54	54	102	57	105	60	110						
33	172	88	51	53	102	59	98								
34	141	79	56	60	83	64	81								
35*	136	83	61	34	71	39	67								
36	133	59	44	29	80	37	88								
37	128	13	10	19	99	24	77								
38	125	57	46	27	101	30	100								
39	113	70	62	64	65	70	77								
40	112	48	43	18	108	22	103								
41	110	51	46	66	76	71	90								
42	98	63	64	40	124	43	116								
43	95	52	55	31	92	36	88								
44	93	56	60	37	163	40	155								
45	88	39	44	59	93	64	84								
46	87	62	71	23	92	28	79								
47	77	47	61	61	97	64	94								

*This child was not included in original monograph because of oversight due to admission in the project late in the third year.

preschool by total days residence during the project.[6] Most of the absences from preschool were on account of illness, in which case the child was kept in the cottage with the control children, or in more serious cases, in the hospital Because the preschool was open only five days a week, the maximum amount of attendance possible was 5/7 or 71 per cent. Several preschool children actually spent more days in the control group than in the preschool group (less than 50 per cent attendance). For instance, while child No. 21 (Table 2) was in residence 432 days, he attended preschool only 188 days, leaving him 244 days in the control environment. He thus had only 44 per cent preschool attendance.

The range of residence for the control children was from 110 to 972 days, for the preschool children from 77 to 970 days. Initial *IQ*'s ranged from 57 to 114 for the control children and from 65 to 163 for the preschool children; two of the preschool children were higher in *IQ* than any control child, but they remained in residence only 93 and 98 days respectively, a shorter residence than any control child. They were merely kept in preschool long enough to transact the mechanics of placement in a foster home The mean initial *IQ* of the control children was 83.5, of the preschool children 86.9. The range in initial ages was from 17 to 69 months for the control children and from 18 to 68 months for the preschool children The respective means were 40.0 and 40.3 months. Attendance in preschool was, of course, zero for the control group; it ranged from 13 to 572 days for the preschool group. Percentage of preschool attendance ranged from 10 per cent to 71 per cent.

D ANALYSES OF VARIANCE

In order to answer the question as to whether preschool and control children differed significantly with respect to *IQ*, analyses of variance[7] were carried through for initial *IQ*'s and final *IQ*'s for four residence groupings, 1-199 days, 200-399 days, 400-599 days, and 600 days or more. Thus preschool children were compared with control children of similar amounts of residence Since, as was stated previously, a number of preschool children

---

[6] In the original monograph, percentage attendance was calculated in relation to the number of days of possible preschool attendance (number of days the child was in attendance divided by number of days the preschool was in session). In order to show more definitely the relation between days in preschool versus days spent in the control environment, per cent attendance in the re-analysis was calculated on the basis of calendar days

[7] Detailed discussion of this technique may be found in: Lindquist, E F, Statistical Analysis in Educational Research Boston: Houghton Mifflin, 1940 Pp. xi+266.

TABLE 3

COMPARISONS BETWEEN PRESCHOOL AND CONTROL CHILDREN OF COMPARABLE AMOUNTS OF RESIDENCE

Group	*N*	Mean residence days	Mean percentage preschool attendance	Mean initial age, months	Mean initial *IQ*	Mean final age, months	Mean final *IQ*	Mean change in *IQ*
		Residence 1-199 days						
Control	12*	135	0.0	42.2	91.3	47.7	94.8	3.5
Preschool less than 50 per cent attendance	6	116	38.8	36.3	92.8	41.5	90.3	—2.5
Preschool 50 per cent or more attendance	10	120	59.5	45.7	99.1	50.3	96.5	—2.6
Preschool 50 per cent or more attendance with one child omitted	9	123	59.4	46.7	92.0	51.4	90.0	—2.0
		Residence 200-399 days						
Control	11	292	0.0	51.5	80.4	61.5	82.8	2.4
Preschool less than 50 per cent attendance	1	269	41.0	27	85.0	36	80	—5.0
Preschool 50 per cent or more attendance	9	288	58.9	52.7	78.8	63.7	85.0	6.2
		Residence 400-599 days						
Control	10*	517	0.0	38.4	83.6	56.0	82.6	—1.0
Preschool less than 50 per cent attendance	8	485	41.5	31.5	83.5	48.4	79.8	—3.7
Preschool 50 per cent or more attendance	9	538	55.7	38.1	84.6	57.1	93.0	8.4
		Residence 600 days or more						
Control	11	853	0.0	27.4	78.1	55.3	67.4	—10.7
Preschool 50 per cent or more attendance	4	850	57.0	30.8	78.3	59.5	81.3	3.0

*These numbers differ by one case from the number reported in previous publications (2, 3) This is due to the fact that through an error in calculation of residence, one child with short residence was wrongly placed in the long residence group spent more days in the control group than in preschool, these children were separated from the remainder of the preschool group. The analyses of variance were thus between three groups of similar residence: (*a*) a control group with no preschool attendance, (*b*) a preschool group with more days spent in the control group than in preschool, and (*c*) a preschool group with more days spent in preschool than in the control group

Table 3 presents the following means for each group: days of residence, percentage of preschool attendance, initial age and *IQ*, terminal age and *IQ*, and change in *IQ*. In the residence grouping 1-199 days, there was one preschool child with an exceptionally high *IQ*. His first *IQ* (Kuhlmann) was 163; his second *IQ* (Stanford) three months later was 155, at which time he was placed in an adoptive home Since the inclusion of a child of this level of ability noticeably affected the mean *IQ*'s for this small group, the means for the remainder of the group when this child was omitted are presented also in Table 3

Table *B* gives the obtained values of *F*, with their corresponding levels of significance. Since in the second residence grouping, 200-399 days, there was only one preschool child with less than 50 per cent attendance, he was necessarily omitted from the analyses of variance here It will be seen that the *F* values on initial *IQ* were all small and none of them significant at the 5 per cent level of confidence. It can be concluded, therefore, that preschool and control children with comparable amounts of residence did not differ significantly with respect to initial *IQ*'s

TABLE *B*
ANALYSES OF VARIANCE
VALUES OF *F*

Residence groupings	Initial *IQ*	Level of significance	Final *IQ*	Level of significance
1-199 days				
(a) all subjects	484	n s*	.260	n.s
(b) omitting one preschool child of 163 *IQ*	033	n s.	.512	n s
200-399 days	.117	n s	.538	n s
400-599 days	018	n s	4 441	5 per cent
600 or more days	013	n s	4 521	n.s.

*n s. means not significant at the 5 per cent level of confidence

With respect to final *IQ*'s, preschool children with residence of 1-199 days and 200-399 days did not differ significantly from control children with comparable amounts of residence For the residence grouping 400-599 days, however, the *F* value for the final *IQ*'s was significant at the 5 per cent level of confidence The obtained *F* value was 4.441; to be significant at the 5 per cent level an *F* value of 3.40 is required, and to be significant at the 1 per cent level an *F* value of 5 61 is required

In order to ascertain the nature of the differences between groups within the residence grouping 400-599 days, *t*'s were computed The results are given in Table *C*

TABLE *C*

t VALUES FOR FINAL *IQ*'S OF CHILDREN WITH RESIDENCE 400-599 DAYS

	t	Level of significance
Control vs preschool with less than 50 per cent attendance	61	n.s*
Control vs preschool with 50 per cent or more attendance	2 301	5 per cent
Preschool with less than 50 per cent attendance vs preschool with 50 per cent or more attendance	2 772	2 per cent

*n s means not significant at the 5 per cent level of confidence; 24 degrees of freedom in these comparisons

From inspection of the tabulation, it will be seen that control children with 400-599 days residence did not differ significantly in final *IQ* from children whose preschool attendance was less than 50 per cent. Table 3 shows that these preschool children had spent only 41.5 per cent of their residence days attending preschool, leaving 58 5 per cent of their days in the control group Both the control group and this preschool group made a slight negative change in *IQ*, the change of the control group was —1.0 point, of the preschool group —3 7 points. In contrast to these groups, the preschool children who had 50 per cent or more attendance made a mean gain of 8 4 points. They attended preschool 55 7 per cent of their residence days This preschool group was significantly higher in final *IQ* at the 5 per cent level of confidence than the control children with similar residence, and they were significantly higher in final *IQ* at the 2 per cent level of confidence than the preschool children of similar residence but with low percentages of preschool attendance.

For the residence grouping 600 or more days the difference in final *IQ*'s was not quite significant at the 5 per cent level of confidence, but very close to it The obtained *F* value was 4 521, with the degrees of freedom in-

TABLE 4
COMPARISONS BETWEEN PRESCHOOL AND CONTROL CHILDREN IN COMBINED RESIDENCE GROUPS

Group	*N*	Mean residence days	Mean percentage preschool attendance	Mean initial age, months	Mean initial *IQ*	Mean final age, months	Mean final *IQ*	Mean change in *IQ*
		Residence 1-399 days						
Control	23	210	0.0	46.7	86.0	54.3	89.1	3.1
Preschool less than 50 per cent attendance	7	138	39.1	35.0	91.7	40.6	88.9	—2.8
Preschool 50 per cent or more attendance	19	200	59.2	49.0	89.5	56.6	91.1	1.6
Preschool 50 per cent or more attendance with one child omitted	18	206	59.2	49.7	85.4	57.6	87.5	2.1
		Residence 400 days or more						
Control	21	693	0.0	52.6	80.7	55.6	74.6	—6.1
Preschool less than 50 per cent attendance	8	485	41.5	31.5	83.5	48.4	79.8	—3.7
Preschool 50 per cent or more attendance	13	634	56.1	35.8	82.6	57.8	89.4	6.8

volved here, an *F* value of 4 67 is required for significance at the 5 per cent level of confidence Unfortunately, there were only four preschool children with residence of this length and, although their final mean *IQ* was 13.9 points higher than that of the control group, the variability was such that the difference in means failed to meet the statistical criterion for significance at the 5 per cent level. There were no preschool children with residence of this length who had less than 50 per cent attendance

These data indicate that children who attended preschool fairly regularly over a period of more than a calendar year made better progress intellectually than other children of equal initial *IQ* who attended irregularly or not at all.

Table 4 gives the means on residence, percentage of preschool attendance, age, and *IQ* for preschool and control groups when the two shorter residence groupings were combined into one grouping of 1-399 days, and the two longer residence groupings were combined into one grouping of 400 or more days. Analyses of variance were computed for the combined groups, and the *F* values are given in Table *D*.

TABLE *D*
ANALYSES OF VARIANCE
VALUES OF *F*

Residence groupings	Initial *IQ*	Level of significance	Final *IQ*	Level of significance
1-399 days				
(a) all subjects	407	n s*	083	n.s
(b) omitting one preschool child of 163 *IQ*	500	n s.	.117	n s
400 or more days	162	n s	6.131	1 per cent

*n s means not significant at the 5 per cent level of confidence

The F values on initial *IQ* were small and not significant at the 5 per cent level of confidence. It can be concluded, therefore, that preschool and control children of comparable residence did not differ significantly with respect to initial *IQ*.

On final *IQ*, the *F* value for the shorter residence grouping was small and not significant at the 5 per cent level of confidence. It can be concluded, therefore, that preschool and control children with residence less than 400 days did not differ significantly in *IQ* change.

For the combined longer residence groups, however, the *F* value on final

IQ was 6.131, significant at the 1 per cent level of confidence. Table *E* presents the *t* values between these three groups.

TABLE *E*

t VALUES FOR FINAL *IQ*'S OF CHILDREN WITH RESIDENCE 400 DAYS OR MORE

	t	Level of significance
Control vs. preschool with less than 50 per cent attendance	1.036	n.s.*
Control vs. preschool with 50 per cent or more attendance	3.514	1 per cent
Preschool with less than 50 per cent attendance vs. preschool with 50 per cent or more attendance	1.800	n.s.

*n.s. means not significant at the 5 per cent level of confidence.

The preschool children with less than 50 per cent attendance occupied a position intermediate between the control children and the preschool children with better attendance with respect to final *IQ* and *IQ* change. They did not differ significantly from either group in final *IQ*; to be significant at the 5 per cent level of confidence, *t* should equal 1.96. The preschool children with more than 50 per cent attendance did differ significantly from the control group at the 1 per cent level of confidence. They made a mean gain of 6.8 points in *IQ*, as contrasted with a mean loss of —6.1 points by the control children, and they stood 14.8 points higher in final *IQ* than the control children.

It can be concluded, therefore, that preschool children in residence more than a year and with fairly regular preschool attendance made significantly better progress in intelligence than other children of equal initial *IQ* in residence for the same length of time who did not attend preschool.

From inspection of Tables 3 and 4 it will be noted that there were some differences between the groups in age. Within each residence grouping, however, the preschool children with 50 per cent or more attendance were quite similar to the control children, and there was no apparent relationship between mean age and mean change in *IQ*. The question of the effect of age is discussed more fully in the next section on correlations.

E. CORRELATIONS

1. *Correlations with Residence*

Correlations of *IQ* with length of residence were computed for the entire preschool and control groups, with the exception that the child of 163 *IQ*

TABLE *F*

CORRELATIONS ON RESIDENCE

			Children 30 months of age or older	
	Control $N=44$	Preschool $N=46$	Control $N=31$	Preschool $N=29$
Days residence vs initial *IQ*	—23	—27	—41	—38
Days residence vs final *IQ*	—66	—13	—.55	—14
Initial *IQ* vs final *IQ*	.48	.72	62	.82
Days residence vs final *IQ*, with initial *IQ* constant	—64	.10	—42	31
Days residence and initial *IQ* vs final *IQ*	70	72	70	83

was omitted from the preschool group. The results are shown in Table *F*

Both the control and preschool groups showed a small negative relationship between initial *IQ* and length of residence (i e., there was a tendency for the children with lower intelligence to remain in residence longer) The correlations were not, however, significant at the 5 per cent level of confidence, and there was no difference between preschool and control groups (r's = —23 and —.27) For the control group there was a fairly high negative relationship (—.66 significant at the 1 per cent level of confidence) between final *IQ* and days residence, indicating that there was a definite tendency for the control children who remained in residence a longer time to be lower in final *IQ* than those who were in residence a shorter time. No such tendency was shown in the preschool group, where the correlation between days residence and final *IQ* was —.13. The difference between the two correlations with final *IQ* for the control and preschool groups (—.66 and —13) was significant at the 1 per cent level of confidence.

In order to ascertain whether there was a significant change within the control group from initial to final test in the correlation with residence, t was computed according to the formula presented by Hotelling[8] for testing the significance of difference between correlations in which one of the variables is identical ($r_{12} - r_{13}$). The identical variable here was days residence, the number of days residence for each child remaining the same in the correlation with final *IQ* as in the correlation with initial *IQ*. The re-

---

[8] Hotelling, H The selection of variates for use in prediction with some comments on the general problem of nuisance parameters. *Ann. Math Stat.*, 1940, **11**, 271-283 The formula may be written

$$t = \frac{(r_{12} - r_{13})\sqrt{N-3}\sqrt{1+r_{23}}}{\sqrt{2}\sqrt{1 - r^2_{23} - r^2_{12} - r^2_{13} + 2r_{12}r_{13}r_{23}}}$$

sultant *t* for the control group was 3.631, significant at the 1 per cent level of confidence. The increase in the size of the negative correlation (from —.23 to —.66) can therefore be interpreted as indicating that the *IQ*'s of the control children had changed significantly in the direction that those children who remained in residence longer became lowest in *IQ*. A similar test applied to the preschool group resulted in a *t* of 1.278, not significant at the 5 per cent level of confidence (Table *G*).

TABLE *G*

t VALUES BETWEEN CORRELATIONS WITHIN SAME GROUPS

	Control *N*=44		Preschool *N*=46		Children 30 months of age or older Control *N*=31		Preschool *N*=29	
	r	*t*	*r*	*t*	*r*	*t*	*r*	*t*
Days residence vs initial *IQ*	—.23		—.27		—.41		—.38	
		3.631		1.278		1.023		2.332
Days residence vs final *IQ*	—.66		—.13		—.55		—.14	
Value of *t* at 5 per cent level		1.960		1.960		2.048		2.056
Value of *t* at 1 per cent level		2.576		2.576		2.763		2.779

From these various analyses it can be concluded that the control children became lower in *IQ* with increasing residence whereas the preschool children did not.

The preschool children showed a tendency to remain more stable in *IQ* than the control children; for the preschool children the correlation between initial and final *IQ* was .72, whereas for the control children the correlation was .48 This difference was not significant, however, at the 5 per cent level of confidence. The obtained *t* was 1.761; for significance at the 10 per cent level of confidence a *t* of 1.6449 is required, and for significance at the 5 per cent level of confidence a *t* of 1.960 is required. The obtained *t*, significant at the 10 per cent level of confidence, is large enough to offer strong presumption that the two regressions differed. The mean initial *IQ* of the preschool children was 85.2, of the control children 83.5 The respective initial ages were 40.4 and 40.0 months.

For the control children the partial correlation of days residence with final *IQ* holding initial *IQ* constant was —.64, while for the preschool children it was positive, .10. The difference between these partial correlations was significant at the 1 per cent level of confidence. These correlations indicate that the control children who remained in residence a longer time were lower in final *IQ* than other control children of equal initial intelligence who had shorter residence. The preschool children did not show this trend

Combining days residence with initial *IQ* in a multiple correlation with final *IQ* resulted in a correlation of .70 for the control children, a considerable increase over the 48 between initial *IQ* alone and final *IQ*. For the preschool children the multiple correlation was the same (.72) as the correlation between initial *IQ* alone and final *IQ*; for them length of residence did not affect the relationship between initial and final *IQ*.

2. *Effect of Age*

It will be recalled that the range of ages at time of initial test was from 17 months to 69 months for the control children and 18 months to 68 months for the preschool children. Because considerable skepticism has been expressed in some quarters in regard to the validity of *IQ*'s obtained at the very young ages, particularly around 18 months, it was decided to compute correlations in which the younger children were omitted. There were 13 control children and 17 preschool children whose ages fell between 17 and 29 months, leaving 31 control children and 29 preschool children who were 30 months of age or older on initial test. The correlations with residence for the latter groups are presented in preceding tables.

The pattern of correlations for the older children is similar to that for the total groups. Initial *IQ* showed a somewhat greater negative relation to amount of residence for the older children, the correlations of —.41 and —.38 were significant at the 5 per cent level, but, as before, there was no difference between control and preschool children. As before, the control children showed a greater negative correlation of residence with final *IQ* than of residence with initial *IQ* (— 55, significant at the 1 per cent level, as compared with — 41, significant at the 5 per cent level), whereas the preschool children showed a smaller negative relation of residence with final *IQ* than with initial *IQ* (—.38 reduced to —.14). The difference between the two correlations of residence with final *IQ* (—.55 for the control children and — 14 for the preschool children) was not significant at the 5 per cent level, however.

When the test for significance of difference between correlations involving the same group was applied to the older control group (correlations — 41 and —.55), *t* was found to be 1.023, not significant at the 5 per cent level of confidence. In the case of the older control children, therefore, the tendency for longer residence to depress the *IQ* was not a significant one.

In the case of the older preschool children, *t* between the correlations —.38 and — 14 was 2 332, significant at the 5 per cent level of confidence, indicating that there was a significant tendency for those with longer residence to take a position within the group equal to those with shorter residence

The correlations of initial *IQ* with final *IQ* were somewhat higher for the older portion of the groups (.62 for control and .82 for preschool) than for the total groups (.48 for control and .72 for preschool). The older preschool children tended to be more stable in *IQ* than the older control children (82 vs .62), but the difference was not significant at the 5 per cent level of confidence, *t* being 1.378 Mean initial *IQ*'s were 82.9 for the older preschool children and 81.6 for the older control children

When initial *IQ* was held constant, the relation of residence with final *IQ* for the older control children was — 42 (significant at the 5 per cent level), while the corresponding correlation for the older preschool children was positive (31) but not significant at the 5 per cent level. The difference between the partial correlations for the older control children and the older preschool children (—.42 and 31) was significant at the 1 per cent level Thus, within the older portion of the total group, control children of longer residence were lower in final *IQ* than other control children of equal initial *IQ* with shorter residence. Such was not the case for preschool children who showed a tendency (not significant) for those with longer residence to be higher in final *IQ* than others of equal initial *IQ* but shorter residence

Combining residence and initial *IQ* into a multiple correlation with final *IQ* resulted in a correlation of .70 for the older control children, representing some increase over the correlation between initial *IQ* alone and final *IQ* (.62). For the older preschool children the addition of residence to initial *IQ* did not increase the correlation with final *IQ*, the respective correlations being 82 and .83.

3. *Conclusion*

The general conclusion from the comparisons of correlations on residence is that there were significant differences in trend of *IQ* change between control and preschool children, even when the preschool children were treated as an intact group without regard to internal differences in amount and percentage of preschool attendance. Since significant differences between control and preschool children were obtained when the computations were restricted to children who were 30 months of age or older, it can be concluded that the presence of children younger than 30 months in the total group was not the crucial factor explaining the differences.

4 *Correlations with Preschool Attendance*

Within the preschool group, correlations were computed between *IQ* and amount and consistency of preschool attendance. These correlations, for the total preschool group (with the one child of 163 *IQ* omitted) and for the children 30 months of age or older, are presented in Table *H*

TABLE *H*
CORRELATIONS ON PRESCHOOL ATTENDANCE

	Total preschool *N*=46	Preschool children 30 months of age or older *N*=29
Days attendance vs initial *IQ*	—29	—40
Days attendance vs. final *IQ*	—07	—15
Initial *IQ* vs final *IQ*	.72	82
Days attendance vs final *IQ*, with initial *IQ* constant	22	33
Per cent attendance vs initial *IQ*	—.14	—12
Per cent attendance vs final *IQ*	19	—01
Per cent attendance vs final *IQ*, with initial *IQ* constant	42	.16
Days attendance and initial *IQ* vs final *IQ*	.74	.84
Per cent attendance and initial *IQ* vs final *IQ*	78	83

There was a negative relationship between days of preschool attendance and initial *IQ;* for the total preschool group the correlation of —.29 was significant at the 5 per cent level, and for the older children the correlation of —.40 was also significant at the 5 per cent level. These correlations indicate that there was a tendency for the children with lower initial intelligence to be present more days in preschool than the children who were initially brighter This relationship would be expected from the high correlation between residence and attendance (.93 for the total group and .99 for the older children) Final *IQ* showed a reduction in the negative relationship (reduced from —.29 to —.07 for the total group and —.40 to —.15 for the older children) Neither correlation with final *IQ* was significant at the 5 per cent level.

When the *t* test for significance of change in correlation was applied it was found that the reduction in negative correlation with days attendance from —.29 to — 07 for the total preschool group was significant at the 5 per cent level of confidence The obtained *t* was 2 060. The reduction in negative correlation from —.40 to —.15 for the older preschool children was also significant at the 5 per cent level of confidence and nearly significant at the 2 per cent level. The obtained *t* was 2.468, to be significant at the 2 per cent level a *t* value of 2.479 is required. It can be concluded, therefore, that the preschool children with the greater amounts of preschool attendance had significantly changed their *IQ* positions towards a higher status within the group

When initial *IQ* was held constant, the correlations of preschool attendance with final *IQ* were positive (.22 and .33) but not significant at the 5 per cent level From these partial correlations it cannot be said that children with greater preschool attendance were higher in final *IQ* than children of equal initial *IQ* and less attendance.

Percentage of preschool attendance was not significantly related to initial *IQ* (—.14 and —.12), nor to final *IQ* (.19 and —.01). The partial correlation of per cent attendance with final *IQ* when initial *IQ* was held constant was positive and significant at the 1 per cent level for the total group ($r = .42$), and was positive (.16) but not significant for the older portion of the group. The *t* test for significance of change in correlations reveals that the change in correlation of per cent attendance with *IQ* from —.14 on initial test to .19 on final test for the total group was significant at the 1 per cent level of confidence, *t* being 3.222 These correlations indicate that children with greater regularity of attendance tended to become higher in *IQ*. The corresponding change for the older portion of the group from —.12 to —.01 was not significant, *t* being .952

The difference between the relationships with percentage of attendance for the total group and the older portion of the group was apparently caused by the fact that the children with low percentages of attendance were preponderately younger children The total group therefore had a greater range of percentage. Of the 15 children in the total preschool group who were below 50 per cent attendance, 11 were less than 30 months of age on initial test. The range of per cent attendance for the children 30 months of age or older was from 44 per cent to 64 per cent, while for the children less than 30 months of age it was 10 per cent to 71 per cent; 86 per cent of the older children, but only 67 per cent of the total group, had attendance 50 per cent or better.

For the total group the multiple correlation of initial *IQ* and per cent attendance with final *IQ* represented a slight increase over the correlation of initial with final *IQ* (.78 was compared with .72); the addition of days attendance to initial *IQ* resulted in an increase of only .02 points in the correlation (74 as compared with 72). For the older portion of the group there were changes of only .01 and .02 points

From the correlations within the preschool group, it can be concluded that there were significant changes in relative *IQ* status associated with number of days of preschool attendance and with degree of regularity of attendance The effect of regularity of attendance was more noticeable among the younger children since most of the older children were quite regular in their attendance.

F. SUMMARY AND CONCLUSIONS

The purpose of this study was to present some new analyses of the data collected in an orphanage preschool project on Binet *IQ* changes of children subjected to orphanage and orphanage-preschool environments. The new analyses utilized for each individual only his initial and terminal intelligence tests.

The control group was composed of children of preschool age who lived in "cottages" under the supervision of untrained help, with a minimum of play materials and a minimum of contacts with adults. A modern nursery school staffed with trained teachers was available to the preschool children five days a week for several hours a day

The preschool group consisted of 47 children with orphanage residence after the project started varying from 77 to 970 days and with attendance in the orphanage preschool varying from 13 to 572 days The control group was composed of 44 children with residence after the project started varying from 110 to 972 days, all with zero days of preschool attendance. Initial ages of the preschool children ranged from 18 to 68 months and terminal ages from 22 to 80 months; for the control children the initial ages ranged from 17 to 69 months and the terminal ages from 22 to 82 months. Initial *IQ*'s of the preschool children ranged from 65 to 163 and of the control children from 57 to 114, two of the preschool children were higher in *IQ* than any control child, but they remained in residence a shorter time than any control child. The mean initial *IQ* of the preschool children was 86 9 and of the control children 83 5. The respective mean initial ages were 40 3 and 40.0 months

Within the preschool group there were several children who attended preschool less than 50 per cent of the calendar days and who therefore spent more days in the control environment than in preschool

Analyses of variance on initial *IQ* and on final *IQ* were made between control children, preschool children with less than 50 per cent attendance, and preschool children with 50 per cent or better attendance in each of four residence divisions No significant differences were found within residence divisions with respect to initial *IQ*'s. Preschool and control children in the two shorter residence groups, 1-199 days and 200-399 days, did not differ significantly in final *IQ*. However, in the residence group of 400-599 days there were significant differences in final *IQ*. The preschool children with 50 per cent or better attendance made a mean increase of 8.4 points in *IQ*, and were significantly higher in final *IQ* at the 5 per cent level of confidence than the control children of comparable residence, who lost —1.0 point They were also significantly higher in final *IQ* at the 2 per cent level of confidence than the preschool children with less regular attendance, who made a mean loss of —3.7 points in *IQ*. Control children and preschool children with irregular attendance did not differ significantly. In the longest residence group, 600 or more days, there were only four preschool children; the preschool children gained 3.0 points in *IQ* while the 11 control children of similar residence lost —10 7 points. The difference in final *IQ* was 13 9 points, but due to the small number of cases, this difference was not statistically significant.

Further analyses of variance were made in which the two shorter residence divisions were combined and the two longer residence divisions were combined. Within the combined shorter residence division no significant differences were revealed between preschool and control groups either in initial *IQ* or final *IQ*. Within the combined longer residence division significant differences were obtained. The preschool children with 50 per cent or better attendance made a mean gain of 6.8 points in *IQ*, as contrasted with a mean loss of —6.1 points by the control children; their final *IQ*'s differed by 14.8 points The difference was significant at the 1 per cent level of confidence. The preschool children with irregular attendance occupied a position intermediate between the other two groups and did not differ significantly from either.

It can be concluded from the analyses of variance that preschool children with regular attendance and in residence more than a year made significantly better progress in intelligence than children of equal initial intelligence in residence for a similar period who did not attend preschool.

Correlations of residence with *IQ* showed significantly different trends in control and preschool groups, even when the preschool children were treated as an intact group without regard to internal differences in amount and percentage of preschool attendance The general tendency of the control group was for longer residence to be associated with increasingly lower *IQ*'s Since significant differences in correlations were obtained also between preschool and control groups when the computations were restricted to children who were 30 months or older, it can be concluded that the presence of children younger than 30 months in the total group was not the crucial factor explaining the differences

Within the preschool group correlations were computed between *IQ* and amount and percentage of preschool attendance From these correlations it can be concluded that there were significant changes in relative *IQ* status associated with number of days of preschool attendance and with regularity of preschool attendance. The effect of regularity was more noticeable among the younger children since most of the older children were quite regular in their attendance

The results of the re-analysis support the conclusions of the original monograph. (*a*) that the control environment produced substantial losses in *IQ* when experienced for long periods, (*b*) that preschool education supplementing the control environment counteracted such losses, (*c*) that whether or not the preschool environment produced gains depended upon the amount and consistency of preschool attendance

References

1 Barrett, H E, & Koch, H L The effect of nursery-school training upon the mental-test performance of a group of orphanage children. *J. Genet. Psychol*, 1930, **37**, 102-122

2 Skeels, H M, Updegraff, R, Wellman, B. L, & Williams, H M. A study of environmental stimulation an orphanage preschool project. *Univ. Iowa Stud., Stud Child Welf.*, 1938, **15**, No. 4 Pp 191

3. Wellman, B L, Skeels, H. M., & Skodak, M. Review of McNemar's critical examination of Iowa studies *Psychol Bull*, 1940, **37**, 93-111.

Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

The Journal of Genetic Psychology, 1944, 65, 265-273.

A PRELIMINARY STUDY OF AUDITION IN THE WHITE RAT*

Department of Psychology, University of Arizona

ALLEN T DITTMANN[1]

The present experiment was designed for use as a basis for the study of the rôle of certain problems of audition in the adaptive behavior of the white rat. It provides not only a means of determining how accurately the rat is able to localize sound but also a simple technique which might be used for training any animal to go toward noise to get food It was patterned after an experiment by Barber (1), which has received but scant attention in reports of later workers—Muenzinger and Gentry (5), Thuma (7), Pennington (6), and Henry (2)—probably because their work followed more closely the technique used by Hunter (3) in a more extensive series of experiments. The findings of the present experiment relative to sound localization *per se* however, were never intended to constitute the end of the work, the main purpose having been the laying of groundwork for later experimentation in audition in animals, a field which has not yet been too thoroughly investigated. Also at least one unforeseen problem which warrants further investigation, that of orientation of the animal with the experimental situation, has arisen in the course of the work

Barber used an eight-sided box in her experiment, much like that shown in Figure 1. The sides of the box were marked off into sections so that there were 32 side-sections around the perimeter of the box, behind any one of which the stimulus could be produced Food was placed on the top edge of the side at the stimulus section, forcing the animals to stand up to take the food, thereby indicating more accurately the side-section where the sound had been localized

The stimulus used most often in Barber's experiment was the tapping of a pencil on a ring of lath-sized wood placed around the outside of the box and supported separately She also used, as stimulus, taps on a tuning-fork resonator, taps on the tuning fork itself, and the sound from an

---

*Received in the Editorial Office on March 2, 1943.

[1] The writer wishes to express his indebtedness to Dr. M R Schneck, the late Dr W D Ellis, and especially to Dr. C L Vaughn for the time, help, and invaluable suggestions of which they so generously gave all during the working out of the present experiment and writing up of this report

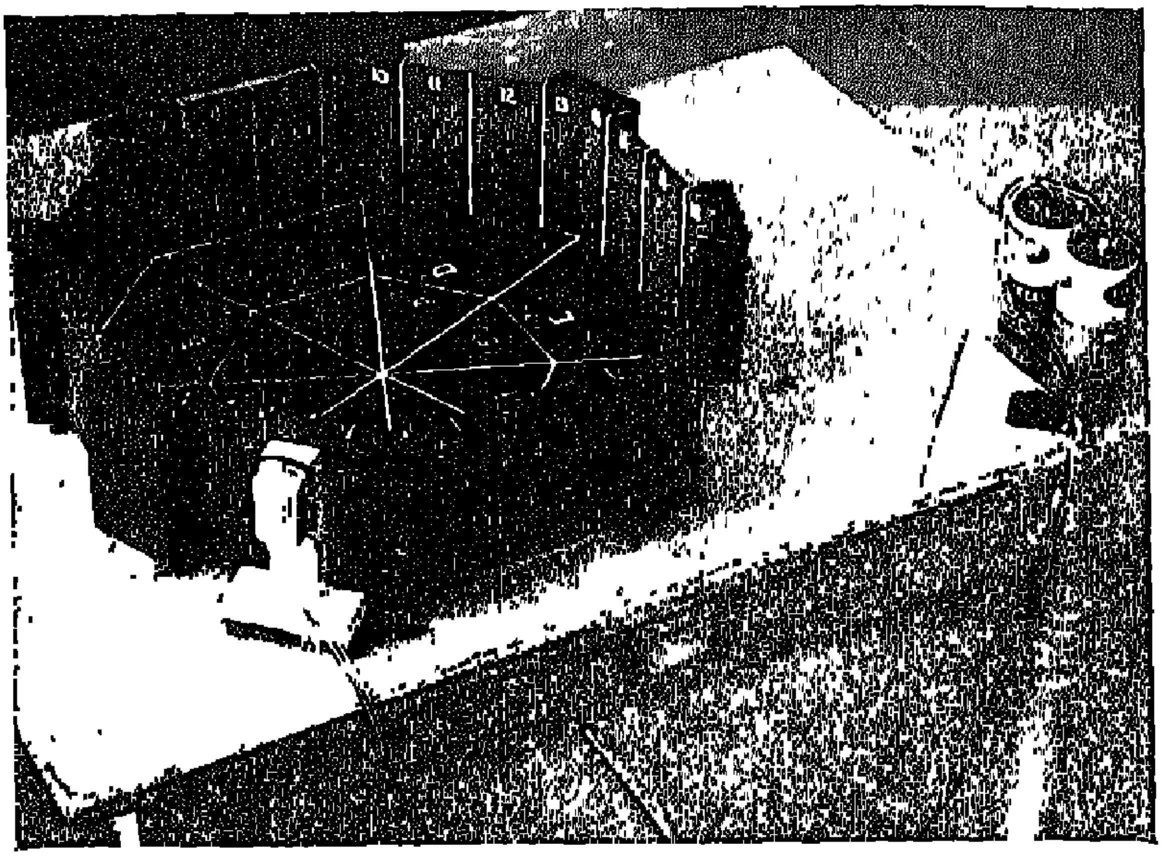


FIGURE 1
The Apparatus with Clicker and Food Arranged for a Trial with Side Section No 23 as the Stimulus Section

organ pipe She found that the animals responded accurately and very quickly, that is, in from 40 to 136 trials, to noise stimuli, but that they failed to respond to pure tones or nearly pure tones

Apparatus and Procedure

In Figure 1 a photograph of the apparatus used in the present experiment is shown. The box was of the same diameter as Barber's, 20 inches The height of the sides was 5 5/8 inches as contrasted with 7 7/8 inches in Barber's apparatus, since preliminary work showed that the animals used here, being rather young at the beginning of the experiment, could not easily stand up very high The sides, made of 3-ply wood, were connected by columns of 2 by 4 inch yellow pine into which *V's* of 135 degrees had been cut to form corners. The floor was of 1 inch white pine in two sections, each 10 inches wide, dowelled and glued together The whole apparatus was painted black There were 24 side-sections marked off instead of 32 as Barber used, since the animals in Barber's experiment were not capable of correctly localizing one out of 32 sound spots, and since it was originally thought that scoring would be made easier with fewer sections Sixteen floor sections were also painted and lettered, to provide a means of recording the animal's route from the start of the trial until the *food had eventually been found.*

The stimulus used in the present experiment was the sound of a buzzer shorted off so that it would produce a click This mechanism was mounted on a wood stand four inches high, and operated by batteries in series with a door-bell push button which always remained in the same place on the table. A good pinna reflex indicated that the sound was easy for the animals to hear. Sponge rubber cushions were cemented to the bottom of the stand to eliminate possible tactual cues from vibrations which might come through the table and floor of the box

The incentive used throughout the experiment was food, in this case small pellets of bran, dried milk, and meat scraps, mashed flat, and about as wide as the 3-ply edge of the sides of the box. One pellet was placed in the stimulus section, and one in each of the adjoining sections, so that the animals would also have reward for trials very nearly correct An attempt was made to keep the motivation strong by regulating the animals' *regular feedings so that they would lose weight* slightly, about 1/2 gram a day, during the course of the experiment Those which lost more than two grams were isolated and fed additional pellets.

Originally five female albino rats, all 14 weeks old at the beginning of the experiment, and all but one litter mates, were to be used as subjects One animal, however, had to be discarded very early in the experiment, because it would not take the food pellets offered as reward for the trials.

Before the experiment proper was begun there was a preliminary period during which the animals were allowed to become familiar with the apparatus For five days the animals were placed in the box and left for increasing lengths of time, two hours the first two days, and five hours the last three On all five days they were fed in the box.

During the first nine days of the experiment the procedure was somewhat different from that of the remainder of the time At first an animal *was placed in the center* of the floor of the box and the stimulus begun immediately The animal's course of travel was then traced by noting down the floor-sections over which it went and the side-sections where it stood up during the entire trial The alteration in procedure came about as a result of the tendency of the animals to go directly to that side of the box toward which their heads were pointed when they were released, to run around nervously for a few seconds, and to stand up at several side-sections in rapid succession before choosing the correct one. The experimenter thought that this short period of nervousness might have been caused mainly by his handling the animals between trials and by the sudden change on being placed in the apparatus, and that all of this might have been aggravated by the sound of the stimulus On the tenth day, then, animal No. 3 was allowed on its fourth trial to go to the side of the box, put its paws up on the edge, to calm down after being handled by the experimenter, to "become oriented" within the apparatus, before the stimulus was begun The animal went almost directly toward the sound Continuing in this manner it was found that the response of the other animals was the same, and consequently the procedure was permanently changed.

Four successive trials a day were given each animal, after which it was fed in a feeding cage. The sequence of side-sections behind which the stimulus was to be sounded was the same for all animals, and was arranged in random order, care being taken, however, that no one side-section be the stimulus section more than once in four or five days. Before each animal's trials for the day the clicker and food were arranged for the first trial. An animal was taken from the home cage, placed in the center of the box, and allowed to become oriented for 5 or 10 seconds (more active animals required more time, while those which appeared frightened needed almost none). The stimulus was then given, at a rate of about one click per second, and the animals route from that time until the food was finally found was noted down, no matter how many errors were made in the meantime. After it had found the food and eaten it, the rat was picked up and placed on the experimenter's arm while he placed the food and clicker in position for the next trial

The trials were scored in terms of per cent error. Since there were 24 side-sections to the box, arranged roughly in a circle, if the first side-section at which the rat stood up were diametrically opposite the correct one, that trial would be twelve sections off, or 100 per cent in error. For example, if the clicker were put behind section No. 23 (as it is in Figure 1), and the first side-section at which the animal stood up after the stimulus was begun were No 11, that trial would be counted 100 per cent error Therefore, if the animal were one section off the error would be 1/12 of 100 per cent, or about 8 per cent, and so forth

RESULTS AND CONCLUSIONS

The accompanying graphs, Figures 2 and 3, give the results in terms of per cent error. Figure 2 shows the average of the trials of all four animals for each day; the curve has been arithmetically smoothed (by averaging for any given day the percentages of that day plus those of the previous and subsequent days, weighting the original day by doubling its percentage) to show any general tendencies Figure 2 shows the individual learning curves of each of the four animals

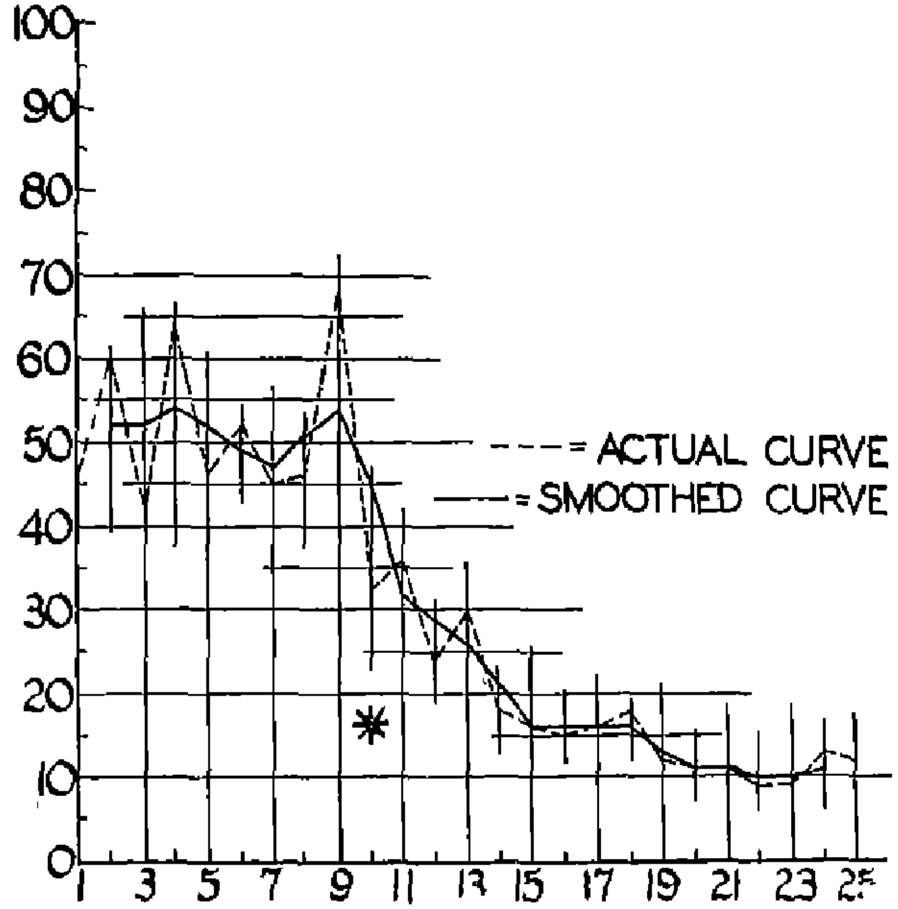


FIGURE 2

THE AVERAGE OF THE FOUR TRIALS EACH OF FOUR ANIMALS FOR EACH DAY IN TERMS OF PER CENT ERROR (The ordinate refers to per cent error, and the abscissa refers to the days. *—indicates first day of change of technique.)

Actually there are twice as many curves in each of these figures as are represented, since the curves for the first nine days of the experiment are quite different from those of the subsequent days There are several factors to which this abrupt change may be attributed. The most obvious explanation is the concurrent change in procedure, in which "orientation"

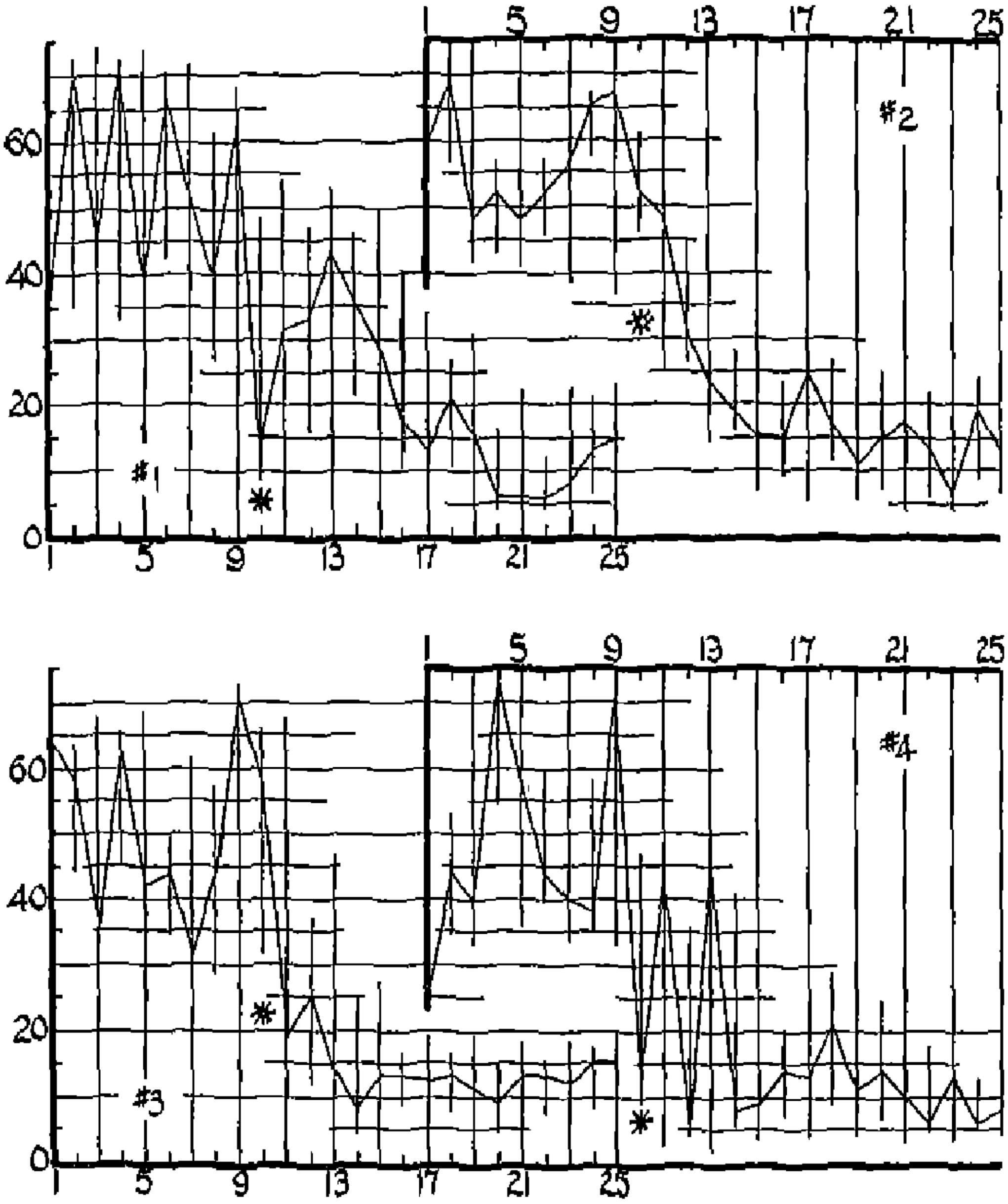


FIGURE 3
Individual Learning Curves of the Four Animals (The ordinates refer to per cent error, and the abscissae refer to days *—indicates the first day of change of technique)

was introduced. But it might be said that such a drop in the curves could have occurred even if the procedure had not been changed Further motivation might have been progressively enhanced by the steady decrease in weight of the animals Only further experimentation can indicate conclusively which of these explanations is the correct one Still it seems highly unlikely that the changed conditions of the experimental technique, and consequently of the entire situation, had nothing to do with the sudden and permanent drop in the learning curves. The animals furthermore showed an immediate change in behavior, which may be most easily expressed in terms of the time they took for the trials Whereas before the tenth day it took as much as an hour and a quarter to run four rats through only four trials apiece, after the tenth day the time was never more than 45 minutes, often as little as 35

To exclude the possibility that the animals were responding to the sight or smell of the food on the edge of the box rather than to the sound of the clicker, a simple control was run On the 22nd and 24th days of the experiment each animal was given two trials (the 2nd and 4th) with no food on the edge of the box After the localization the animals were immediately picked up and given reward in the experimenter's hand On these control trials it was found that the average error was only 10 per cent On the 22nd day the percentages of error were 10 with the food, 8 without, on the 24th day they were 14 with food and 13 without This control demonstrates that neither sight nor smell, nor any combination of these, was essential to the accuracy of the performance

From the results shown above and in Figures 2 and 3 it may be concluded that in a situation similar to Barber's (1) the white rat is able to localize the sound of a click with an error of something less than 13 per cent. Individual learning curves presented (Figure 3) indicate that the individual animals' learning was not basically different from that of the average of all of them This accuracy may be attained in as few as 48 trials by some animals Owing to the change in procedure of the tenth day, which was immediately followed by far more accurate performance, it is felt that far fewer trials would have been required if the altered procedure had been used from the beginning of the experiment. This would seem to support one of Thuma's (7) findings, that while in one situation no results could be elicited, in another situation, outwardly not greatly different, the results were positive. In a tone-silence discrimination there was no indication of learning until the animals were forced by a slight change in the apparatus to hesitate a moment before making the choice

SUGGESTIONS FOR FURTHER WORK

1. First a few mechanical changes in the apparatus have suggested themselves·

a A release cage in the center of the box would force the animals to hesitate longer before starting for the food, and might thereby increase the learning speed.

b The sides of the box might be slightly higher, making the animal stand straighter while getting the food, and making it possible to tell more exactly where the sound had been localized. If more accuracy could be provided in this way more side sections could be added, increasing the fineness and reliability of the scoring.

2. The stimulus used may not have been one which was very natural to the animals. Katz (4, pp. 7-110) reports experiments in sound localization in which a cageful of live, active mice had to be used as stimulus for cats, clucking mother hens for baby chicks, etc., since these animals would not respond to common laboratory sound producers as buzzers and clickers. It is conceivable that with some other stimulus rats could be made to demonstrate their abilities in sound localization to better advantage than the above results might seem to indicate. It may also be that in larger apparatus, even an out of doors plot of grass 20 feet wide instead of 20 inches, greater accuracy might be brought to light.

3. The whole question of "orientation," or whatever psychological phenomenon it was which altered the course of the learning curves so significantly, has been at most posed here. Further experimentation is necessary in which the second procedure of the present experiment is used from the start Controls could then be exercised in which the original procedure is introduced later in the experiment, with some animals continuing on with the "orientation" procedure. Other trials should be run in which the beginning procedure of the present experiment is used throughout in order to discover if learning could ever be evidenced by this method. For these, as well as for any further experiments, of course, it is imperative that greater numbers of animals be used if quantitative results of any reliability are desired

REFERENCES

1 BARBER, A. G. The localization of sound in the white rat. *J Anim. Behav.*, 1915, **5**, 292-311

2 HENRY, F M. Audition in the white rat: III Absolute and relative intensity thresholds *J Comp. Psychol.*, 1938, **22**, 45-62

3. HUNTER, W. S The auditory sensitivity of the white rat. *J Anim Behav*, 1914, **4**, 216-222

4. KATZ, D. Animals and Men. Studies in Comparative Psychology. London: Longmans, Green, 1937.

5 Muenzinger, K F, & Gentry, E. Tone discrimination in the white rat *J Comp. Psychol*, 1931, **12**, 195-206.

6 PENNINGTON, L. A The auditory localizing behavior of the white rat in relation to cerebral function I. Preliminary study *J Genet Psychol.*, 1935, **46**, 264-283

7. THUMA, B. D The response of the white rat to tonal stimuli. *J. Comp Psychol.*, 1932, **13**, 57-86.

Department of Psychology
University of Arizona
Tucson, Arizona

The Journal of Genetic Psychology, 1944, 65, 275-279.

THE INHERITANCE OF SHYNESS IN DOGS*

Brandon State School, Brandon, Vermont

FREDERICK C THORNE, M D.

In evaluating the results from experiments on animal behavior, it is particularly important to study the psychological characteristics of the experimental animals in order to determine whether they constitute a representative sample of the species as a whole This is particularly important in the case of laboratory animals which are obtained from the city pound or bought cheaply because of some outstanding imperfection and which are therefore utilized in animal experiments because of their unsuitability for anything else.

In a previous experiment (1) on approach and withdrawal behavior in a group of 300 dogs, studied at the Cornell University Experimental Morphology Station,[1] it was discovered that about 25 per cent of the animals showed varying degrees of unfriendliness which was not modified by attempts at taming carried out over a 30-day experimental period These animals showed a persistent fear response which was interpreted as being hereditary, since it was observed soon after birth before opportunity, for learning and also because it was not modified by a variety of incentives to be friendly

Since publication of the original data, information has become available concerning the breed and genetic relationships of the animals studied in the experiment. The experimental animals consisted of (*a*) a group of purebred dogs purchased from various sources for the purpose of establishing groups representative of the various morphological types, and (*b*) groups of hybrid animals obtained by crossing various morphological types among the pure-bred dogs The classification of breeds studied in this experiment is presented in Table 1.

The original experimental group consisted of 181 dogs, but three animals died and are not included in the final report Eighty-two or 46 per cent of the 178 experimental animals showed consistent withdrawal behavior which

---

*Accepted for publication by C J. Warden of the Editorial Board, and received in the Editorial Office on March 11, 1943.

[1] Acknowledgment is given to the Department of Anatomy, Cornell Medical College, for permitting study and publication of the data gathered by the late Dr. Charles R. Stockard

TABLE 1

SHOWING THE BREED AND HYBRID TYPE OF 181 ANIMALS INVESTIGATED IN AN EXPERIMENT ON SHYNESS IN DOGS AT THE CORNELL UNIVERSITY EXPERIMENTAL MORPHOLOGY STATION

Pure bred types		Hybrid types	
Bassett Hound	4 animals	Bassett-Shepherd	15 animals
Bloodhound	2	Bassett-E. Bull	19
Boston Bull	2	Bassett-Dachshund	15
English Bull	3	Bassett-Salouki	11
Cocker Spaniel	3	Dachshund-Brussels	7
Dachshund	5	Dachshund-B. Bull	11
Great Dane	1	Dachshund-Peke	2
Labrador Husky	1	Dachshund-Cocker	11
Pekinese	4	Gt Dane-St Bernard	17
Pomeranian	4	Gt Dane-Bloodhound	1
St. Bernard	1	Labrador-Chow	3
Salouki	2	G. Shepherd-E Bull	1
English Setter	2	Salouki-Peke	2
German Shepherd	7	Pomeranian-Peke	3
		Bull Terrier-Bassett	3
		Salouki-Bassett × Salouki f_1	6
		Peke-Pomeranian × Peke f_1	3
		Bassett-Bassett × Shepherd f_1	1
		E. Bull-E. Bull × Shepherd f_1	8
		E. Bull-E Bull × Shepherd f_1	2

was unmodified by incentives to taming. Of the 82 shy animals, genetic analysis revealed that 43 or 52 per cent were descendants of a single Bassett bitch named Paula who was extremely shy and known as a bad "fear-biter." In spite of her extremely fearful and generally undesirable social behavior, this animal was extensively crossbred with other morphological types because of her great fecundity.

An analysis of the genetic records indicates that this Bassett bitch was successively mated with a Salouki, a Dachshund, an English bull, and a German Shepherd. In each case the male animals were considered to be normal friendly dogs who had never shown excessive shyness or withdrawal behavior The offspring of these various matings are graphically tabulated in Figure 1 together with notations from the experimental records indicating whether each animal was rated as being friendly or shy. In several instances siblings in the second generation were mated, and in two instances the f_2 animals were mated with unrelated animals of the same breed. It is unfortunate that all members of the various litters could not be included in the chart, but they had been disposed of before the experiment began.

The original litter obtained by crossing the shy Bassett bitch with a

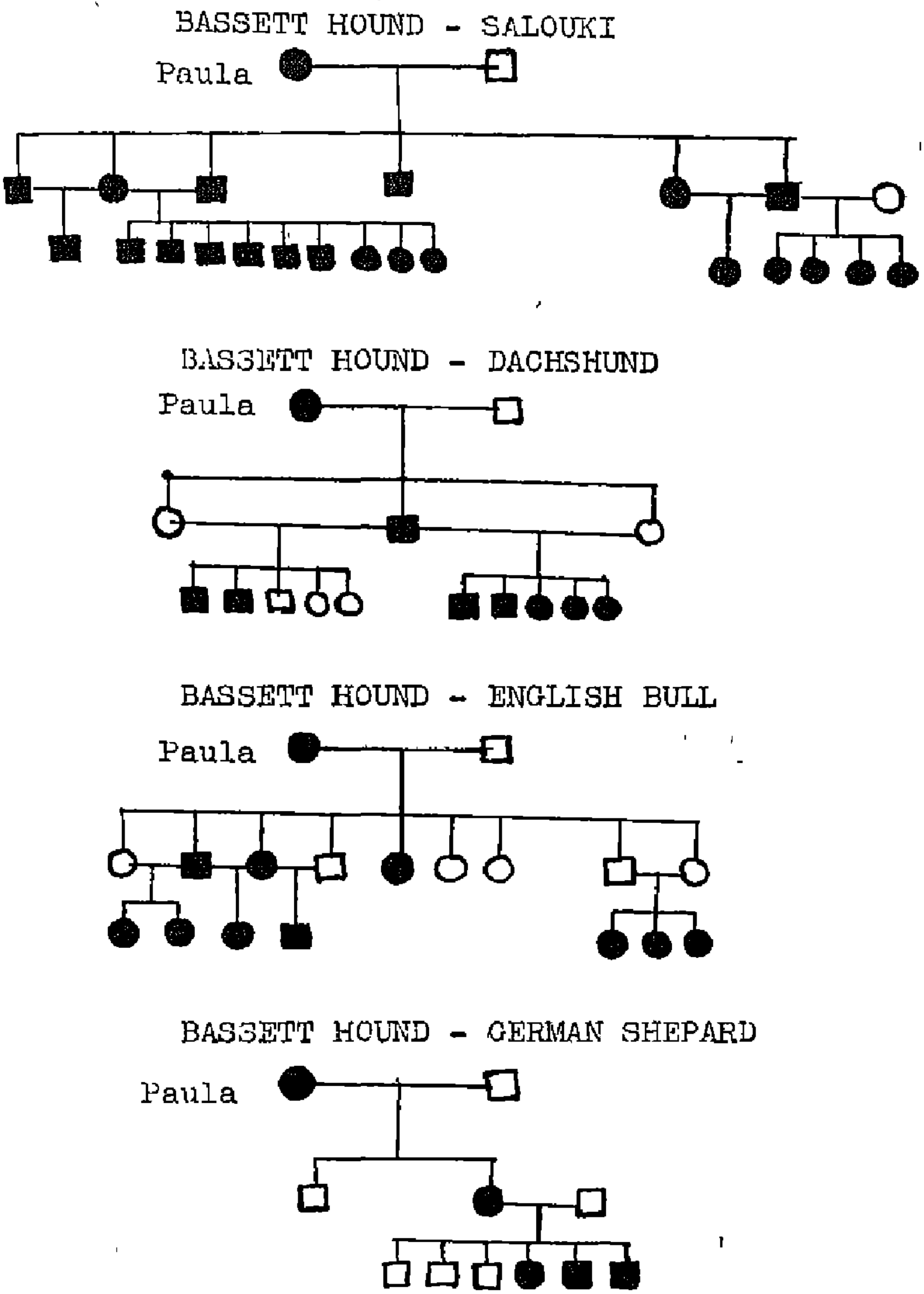

BASSETT HOUND - SALOUKI
Paula
BASSETT HOUND - DACHSHUND
Paula
BASSETT HOUND - ENGLISH BULL
Paula
BASSETT HOUND - GERMAN SHEPARD
Paula

FIGURE 1

Showing the Family Trees Resulting from Successive Matings of a Shy, Fear-Biting Bassett Hound (Paula) with Normal Males of Four Different Pure-Bred Types

(Shy animals are represented in black symbols and friendly animals with white symbols.)

friendly Salouki yielded six pups all of whom became exceedingly shy dogs. Three matings of shy f_1 animals produced 11 pups all of whom were shy Another mating of a shy f_1 male with a friendly female Salouki produced a litter of four all of whom were shy All of the Bassett-Salouki hybrid animals were exceedingly shy, unfriendly, and generally unsatisfactory as pets.

The litter resulting from the Bassett-Dachshund mating contained two friendly females and one shy male. This male was mated with each of the friendly females and 7 out of the 10 pups from the two litters were shy unfriendly dogs

The mating between Paula and a friendly male English Bull resulted in a litter containing two shy females and one shy male out of nine pups. Two matings between shy and friendly f_1 animals resulted in three shy pups. One mating between the shy brother and sister produced an unfriendly pup, while another mating between two friendly dogs resulted in three shy pups The original litter from the Bassett-German Shepherd mating contained one shy female and two friendly males. One mating between the shy female and one of her friendly brothers resulted in a litter of six pups three of whom were shy animals

DISCUSSION

It is now apparent that the high percentage (46 per cent) of shy unfriendly animals encountered in the experimental group of 178 dogs was directly caused by the use of a shy fear-biter (Paula) as one of the principal breeding animals from which the various hybrid strains forming the colony were obtained Fifty-nine descendants of the shy Bassett hound (Paula) were traced and studied and 43 were shy unfriendly animals while 16 were friendly and good-tempered. Although shy animals were encountered in other pure and hybrid types, in no other instance was there such striking evidence of the hereditary transmission of a psychological trait One of the pure-bred Dachshunds was an extremely shy animal and this trait appeared in seven descendants but not in such an impressive degree

It is fortunate that this experimental group of animals was originally bred for the study of the inheritance of morphological characteristics rather than for purely psychological studies. While the presence of an abnormally high percentage of shy animals probably had no appreciable effect on the morphological investigations, this atypical sampling might introduce an important error in studies of physiological traits. It seems important to emphasize that great care should be taken in the selection of brood animals from which the experimental groups are derived

Unfortunately it is not possible to reach definite conclusions concerning the inheritance of shyness in dogs from this study, because of the fact that complete records were not available concerning all the members of the various litters Some members of the litters were assigned to other experiments or had been disposed of before the experiment began so that it was not possible to include all the members of all the litters in the study. No attempt has been made therefore to analyze the data quantitatively The incidence of such a high percentage of shy unfriendly animals among the descendants of Paula suggests that the trait is a dominant characteristic Under normal biological conditions, the extremely shy animal or fear-biter does not survive as effectively as the more friendly animal and this is a possible explanation why such animals do not exist in larger proportion in a normal canine population.

SUMMARY

Analysis of the genetic data concerning an experimental group of 178 dogs, used in a study of approach and withdrawal behavior, reveals that a sampling error was responsible for the abnormally high percentage of shy animals reported in the experiment Of 82 shy animals, 43 or 52 per cent were descendants of an exceedingly shy Bassett hound who was known as a fear-biter. Fifty-nine descendants of this shy dog were traced and 43 or 73 per cent were also shy unfriendly animals. It is suggested that this excessive shyness is caused by the inheritance of a dominant characteristic and is therefore unsusceptible to modification through learning and training

REFERENCE

1. THORNE, F C Approach and withdrawal behavior in dogs *J Genet. Psychol.*, 1940, 56, 265-272

Brandon State School
Brandon, Vermont

The Journal of Genetic Psychology, 1944, 65, 281-289

INFANT LEARNING AS A DEVELOPMENTAL INDEX*

Department of Psychology, Northwestern University

JOHN J. B MORGAN AND SARAH S MORGAN

The experimental work to be reported in this paper is based on the hypothesis that as a child matures he becomes competent to learn tasks of increasing complexity. If this hypothesis is valid, and if it were possible to evolve a series of learning problems which were graded according to complexity, this series could be used as a developmental scale. By means of it one could determine the stage of development of any given child, and it would also be possible to follow a particular child in order to determine whether he was progressing as he should.

This goal of a series of graded learning problems can be reached only by carefully working out each problem which may later find a place on such a scale. We believe we have found one such problem; it cannot normally be learned by an infant who is less than 54 days of age and is normally learned by infants who are older than 65 days Other learning situations are being studied, some harder than this one and some easier. Whether they turn out to give results which are as clear-cut as the one to be reported in this paper remains to be seen. The work is being done at The Cradle in Evanston and we are very grateful to The Cradle Society for permitting us to do this work under conditions which are so ideal.

The lesson which we set for the infant to learn was to react to a specific visual stimulus by winking A puff of air was used as the unconditioned stimulus to produce an eye wink. The conditioned stimulus, to which the infant was to learn to respond, was the movement of the hand, which held the bulb used to produce the air puff, into position in front of the infant's face.

Before beginning the experiment, it was necessary to determine whether the child already had learned to respond to the visual stimulus by winking and whether he would invariably wink when stimulated by the puff of air.

Each infant used in the preliminary experiments was given a series of 50 visual presentations of the hand containing the bulb without any puff. The average number of winks which coincided with the movement of the hand to position before the face was 4 in a series of 50.

---

*Received in the Editorial Office on March 19, 1943.

Each infant used in the preliminary experiments was, likewise, given a series of 50 puffs of air directed into the eye. The puff was delivered by squeezing a baby syringe held in the hand The air did not strike the face at exactly the same spot each time but the puff was given as uniformly as could be done. Every child responded invariably to each puff of air by winking his eyes. The response was not limited to a wink, however. In most cases there were other activities which varied from infant to infant They included gasping, turning the head to one side, retracting the head, waving the arms, kicking the legs, stiffening the body, cooing on the part of some of the older ones, and secondary smiling as the infant relaxed from the tension produced by the stimulation of the air puff

Most of the infants seemed to enjoy the whole procedure, and those who learned the lesson were inclined to smile after each response. One got the impression that they were having a good time. We believe that this last factor is a significant one in infant research. While it is important to use a stimulus which is adequate to produce the desired response, it is also important that the infant be not irritated by it, that is, if learning is the desired goal

Before we began the planned series of experiments, a considerable amount of exploratory work was necessary. At first this consisted of random trials with infants of varying ages From these we gathered the impression, vague as it was, that infants below the age of 45 days did not succeed very well in learning to wink to the visual stimulus alone, and it seemed that infants over 70 days of age learned the lesson very readily. It was decided to begin with some infants at the lower range and to follow them through by daily tests until they did learn.

PRELIMINARY EXPERIMENTS

Eight infants were used for these longitudinal studies. Each was tested over a period of days Some of the series are incomplete and, for that reason, the validity of the work should not stand or fall on the basis of these preliminary results. They are presented in some detail because they are suggestive, they confirm our later results, and they pointed the way to important changes in procedure which we adopted in the main experiment.

The method used was to bring the bulb into position before the infant's face and to produce the puff immediately Ten paired presentations were given in a series and then a test was given, consisting of a visual presentation of the bulb without the puff of air

TABLE 1

Conditioning of Infants of Different Ages to Visual Stimulus*

Infant		Age in days																			
		45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
A	Tests passed			1	1	0	2														
	Series given			10	8	9	9														
	Percentage			10	12	0	22														
B	Tests passed										6	10	6	7	3		15				
	Series given										14	12	6	8	3		15				
	Percentage										43	83	100	87	100		100				
C	Tests passed												9		13	16	4		7		4
	Series given												21		17	17	4		7		5
	Percentage												43		76	94	100		100		80
D	Tests passed											2	2	3		5		1			
	Series given											6	9	9		12		3			
	Percentage											33	22	33		42		33			
E	Tests passed									0	0	1		6							
	Series given									6	9	6		15							
	Percentage									0	0	17		40							
F	Tests passed													0	4	0					
	Series given													8	8	8					
	Percentage													0	50	0					
G	Tests passed										6			45	21						
	Series given										68			113	21						
	Percentage										9			40	100						
H	Tests passed	0	0	0	0			0	0	1	5	4				2	2	14	13		
	Series given	42	41	24	12			85	7	30	65	43				24	20	53	30		
	Percentage	0	0	0	0			0	0	3	8	9				9	10	26	43		

*After each series *of 10 paired presentations* of the hand holding the bulb accompanied by a puff of air, a test was given by presenting the hand containing the bulb without the puff The second line for each infant indicates the number of tests given, and the third line gives the percentage of tests passed

The results of these preliminary experiments are shown in Table 1. For each of the eight infants are given the number of correct tests passed on each day, the number of series given on that day, and the percentage of tests passed. The percentages from this table are shown in scatter form in Table 2

TABLE 2

SCATTERGRAM OF PERCENTAGES FROM TABLE 1

	Ages in days																			
Percentage	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
100												1		2		2		1		
95-99																				
90-94															1					
85-89													1							
80-84											1									1
75-79														1						
70-74																				
65-69																				
60-64																				
55-59																				
50-54														1						
45-59																				
40-44										1		1	2		1			1		
35-39																				
30-34											1		1				1			
25-29																	1			
20-24						1						1								
15-19											1									
10-14			1	1												1				
5-9										2	1				1					
0-4	1	1	1	1	1		1	1	2	1			1		1					

These results suggest that the period between 50 and 60 days of age marks a critical change in the ability of these infants to learn the conditioned wink reaction as set up in this particular experiment The writers also got the impression, which seems to be borne out by the figures, that if a particular child was not mature enough (or incompetent) to learn the task involved in the experiment, extensive practice could not wholly make up for such inability. For example, Infants *B*, *C*, and *G* learned quickly; whereas Infant *H* did not appear to learn at all before the age of 54 days, and not too quickly after that time up to 64 days of age. If practice were sufficient to overcome incompetence, this infant should have learned much more quickly than he did, for he was given 2110 paired presentations before he gave one correct reaction to the test. In all he was given 4760 paired presentations and, even so, was not very superior in his learning. These combined findings suggest that at least three factors operate in producing these results, namely maturity, inherent ability, and practice.

Furthermore, the experimenters derived some suggestions as to procedure which are not revealed by the figures presented, but which came from qualitative observations of the infants. It was noticed that, as the infant learned to wink to the visual presentation of the hand holding the bulb, he would anticipate the puff by various signs, even though the attempt was made to give the puff as soon as the hand came into position before the infant's face. These anticipatory reactions told us clearly that the infant was learning his lesson. Consequently, a few cases (not presented in our tables) were tried in which the interval between the visual presentation and the puff was lengthened This procedure seemed so much more effective that it was used in later experiments.

MAIN EXPERIMENT

In order to determine whether the ability to learn to respond to the visual presentation of the hand containing the bulb was dependent upon maturation, it was decided to test a number of infants, using each infant on only one day, and at such ages that they ranged from a point well below the critical area of 50 to 60 days to another point well above this area. Forty-two infants were used in this series.

In each instance, a series of 10 visual presentations of the hand containing the bulb was given before the experiment in order to determine whether the infant was already conditioned to react to such a stimulus by winking. In a few cases, the first appearance of the hand produced a wink, but succeeding ones failed to do so. That is to say, the results obtained from this preliminary check showed that the conditioning from later training cannot be explained on the basis of a response already established.

The procedure for this experiment was as follows: The bulb was shown on a mental count of one by the examiner, held in position before the infant's face during the count of two, and squeezed on the count of three After a number of infants had been tested the examiner checked his speed of counting with a metronome and found it to be about the rate of 100 per minute. This made the average interval between the visual presentation and the puff 1 8 seconds.

It might be pointed out that this procedure was quite similar to the well-known prompting method of learning. The blub was brought into position, the infant allowed an interval of 1.8 seconds to indicate whether or not he knew what to do (to wink); and whether he winked or not, he was prompted, that is, given the stimulus to make him wink. If he had anticipated the puff by winking, the puff was a sort of corroboration. "Yes, you did the right thing." If he did not anticipate the puff by a wink, he was given the answer—the puff made him wink.

The interval between presentations was about five seconds This was not kept rigid because the infants varied in their readiness to respond, and in the time it took them to recover from the puff Some were placid and made few accessory movements. Others were excitable and it took longer for them to subside after each puff.

The training with each infant was continued until the infant gave 10 successive winks to the visual stimulus before receiving the puff of air, or until he was given over 100 presentations. In computing the results, it was decided to use another criterion for success than a run of 10 conditioned winks without error. It was found that even an infant who learned very quickly might momentarily have his attention diverted and fail to respond. Consequently it was decided to use the following standard: Whenever the infant succeeded in giving eight winks in any 10 successive presentations, the number of the last trial of such a series was used as his score For example (letting *x* indicate failure and *p* indicate a conditioned wink), suppose an infant's record was as follows, the numbers indicating the sequence of trials· 1, 2 *x*, 3 *x*, 4 *p*, 5 *x*, 6 *p*, 7 *p*, 8 *p*, 9 *x*, 10 *p*, 11 *p*, 12 *p*, 13 *p*, 14 *p*, 15 *p*, 16 *p*, 17 *p*, 18 *p*, 19 *p* If we used 10 uninterrupted winks for a criterion, the score for this infant would be 19; whereas we used a criterion of eight successes in a series of 10 giving this child a score of 13 This criterion enabled us to pick successfully the infant who could learn, and avoided erratic scores such as might occur because of some chance disturbance

The results for this series of experiments are given in Table 3 Scores are plotted against age In order to get the pass and fail scores in the last two columns of this table, we set up the standard that an infant must learn the lesson assigned in this experiment in less than 100 presentations

It will be noticed that no infant of 53 days of age or younger passed, whereas only one infant 66 days of age or older failed.

DISCUSSION

From these results, it seems to be clear that maturation is an important factor in determining the ability of an infant to learn the simple conditioned

TABLE 3

SCATTERGRAM OF SCORES MADE BY INFANTS OF DIFFERENT AGES ON LEARNING TEST*

Ages in days	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	Above 100	Total pass	Total fail
120-124																1		1
115-119																		
110-114			1														1	
105-109																		
100-104	1																1	
95-99			1														1	
90-94										1							1	
85-89	1	2															3	
80-84		3															3	
75-79	2	2	1														5	
70-74		1	3														4	
65-69				1		1				1				1		1	4	1
60-64		1														1	1	1
55-59			1					1								3	2	3
50-54								1								5	1	5
45-59																1		1
40-44																1		1
35-39																		
30-34																1		1
25-29																		
20-24																		
15-19																1		1

*The score was determined as follows When the child responded to the visual stimulus before the puff was given for 8 stimulations in any 10 successive paired presentations, his score was the total presentations that had been given. Any infant who did not meet this requirement in 100 presentations was considered to have failed

reaction as set up in this experiment In the normal child, we can infer that the transition from incapacity to ability is between the ages of 54 and 65 days. We should be inclined to say that if an infant is able to meet our criterion before 54 days of age, he thereby shows signs of being precocious; if he is unable to meet our criterion when older than 65 days of age, he thereby shows signs of being retarded

The nature of such precocity or retardation we are not prepared to state. At least three factors could contribute: physiological maturation, sensory acuity, and intelligence. However, even though the causes of individual variability cannot be specified, it seems valuable to have an index of normal progress.

In the case of the one infant who failed to meet our criterion, an interesting group of speculations arose. This child was a perfect physical specimen. She was pretty, had large eyes, an appealing smile, and would have seemed intelligent to any casual observer She was not placed for adoption, but was returned to her natural parents. She was followed for a longer period of time than the other infants and finally, at 127 days of age was able to learn our lesson. Before that time she reacted to our experiment in a manner quite similar to the reactions of much younger infants. When the bulb was brought before her eyes, she fixated on it and opened her eyes wider, but did not wink till the puff was given She worked just like a machine If regularity and uniformity of performance were to be used as a criterion of intelligence, she would have been a genius But such is not the case. The intelligent infant shows variability and changes his performance to such a degree that he can modify his actions with experience, selecting the better responses and eliminating the inferior ones.

In short, we believe that an intelligence factor enters into this situation as well as a maturation factor. We are inclined to believe that, before a certain stage of maturation has been reached, the infant is incompetent to learn the lesson involved in our experiment, but that, when he does not learn it after he is physically mature enough to do so, such failure may be an index of inferior intelligence This is advanced as a hypothesis for which we have inadequate evidence This infant was given a Cattell test at five months of age and, according to it, had an *IQ* of 76 This is suggestive but is not sufficient to settle the point at issue. If a number of deviates, both precocious infants and retarded ones, according to the conditioned wink reaction criterion established by this experiment, are followed for a period of years, we may be able to verify or to repudiate this hypothesis.

SUMMARY

Fifty infants at The Cradle in Evanston were given paired presentations of a visual stimulus and a puff of air in order to determine the age at which they became competent to learn a conditioned wink reaction to the visual stimulation, and to discover other factors than maturation which might contribute to such learning.

There is clear-cut evidence that the normal infant cannot learn this conditioned reaction before the age of 54 days, and that the normal child of over 65 days of age can learn it

There is some evidence to suggest that if the infant is not prepared to learn this conditioned reaction, excessive drill is not very effective in teaching it to him.

There is also the suggestion, still to be tested, that failure to learn this

conditioned reaction, by an infant who is physically mature enough to do so, may be an indication of intellectual retardation or lack of intelligence.

Department of Psychology
Northwestern University
Evanston, Illinois

The Journal of Genetic Psychology, 1944, 65, 291-300

AGE DIFFERENCES IN RELIGIOUS BELIEFS AND PROBLEMS DURING ADOLESCENCE*

School of Education, Syracuse University, and Goucher College

RAYMOND G KUHLEN AND MARTHA ARNOLD

Problems in the realm of values, philosophy of life, and religion have long been considered a major area of adjustment confronting adolescents. Marked changes in religious views are assumed to occur and many crucial problems to arise At the college level it has been fairly well demonstrated that, on the average, profound change does not occur. That seniors are only slightly more liberal in outlook than freshmen has been demonstrated in studies by Katz and Allport (5), Dudycha (3), and others However, adequate evidence indicating age differences in beliefs held and problems faced during the adolescent period is not available for ages below college. Dimock, who has reported one study of younger adolescents, concludes that his data reveal "the relatively static nature of religious ideas of the adolescent during the four-year period (12-16) encompassed by this study," thus apparently contradicting popular opinion (1, 2, p. 168) It is probable, however, that the type of analysis followed by Dimock in his research has obscured a number of trends,[1] and his data can hardly be accepted as evidence demonstrating religious ideas not to change during adolescence

The research summarized in the present paper was an attempt to explore the nature of religious beliefs held during the adolescent period, and to test by at least one type of evidence the hypothesis that adolescence is a period of increasing religious problems. It differs from other investigations in that the results have been analyzed both with respect to specific problems faced and beliefs held, and with respect to relative age periods during adolescence, thus providing an age perspective in which to view such specific changes as might occur

PROCEDURE

A questionnaire was prepared which listed 52 statements representing vari-

---

*Received in the Editorial Office on March 22, 1943

[1] See the "Discussion" section of this paper for a further elaboration of this point Franzblau (4) has shown that beliefs of Jewish adolescents become more liberal with increased age, and MacLean (6) has presented data indicating the percentage of children of adolescent years who hold certain specific beliefs

ous religious beliefs (sample statements are contained in Table 2) which were to be marked according to whether the subject "believed" the statement, did "not believe" the statement, or did not know whether he believed it or not but had "wondered about" it. A second part of the questionnaire listed 18 problems of a religious sort[2] with directions asking the subject to circle an "*N*," an "*S*," or an "*O*" depending upon whether he felt that a particular problem troubled or bothered him "never," "sometimes," or "often."

Responses from 547 sixth, ninth, and twelfth graders were obtained. The exact number of subjects and the mean age of each grade group are contained in Table 1. The significance of the use of these three grade groups in the study of adolescence lies in the fact that the sixth grade group is largely prepubescent, the ninth grade group pubescent (at least for boys; girls mature somewhat earlier), and twelfth grade group post-pubescent. The sampling of these three phases of development during the age range in which adolescence normally occurs should be sufficient to provide evidence of any marked trend occurring during this period. During the discussion to follow, these groups will be referred to as 12, 15, and 18 year-olds. As might be expected, the groups were quite heterogeneous with respect to church membership. Among the 12, 15, and 18 year-old groups respectively there were 19, 33, and 10 Catholic boys and 26, 19, and 18 Catholic girls. The rest were scattered among Protestant denominations, with an occasional child of Jewish faith, or who indicated no church attendance.

TABLE 1

NUMBER OF CASES, MEAN AGE, AND STANDARD DEVIATION IN YEARS OF GROUPS STUDIED

Grade	N	Boys Mn. Age	*SD*	N	Girls Mn. Age	*SD*	Total No.
6	80	12.4	1.2	94	12.0	0.9	174
9	128	15.2	1.0	115	14.9	1.0	243
12	49	18.1	0.8	81	17.8	1.0	130

CHANGES IN RELIGIOUS BELIEFS

The findings were analyzed first by ascertaining what proportion of each age group checked each statement indicating belief, disbelief, or uncertainty regarding that statement. Certain results of this analysis are summarized in Table 2. An examination of these findings indicates clearly that a num-

---

[2] Taken from the Problem Check List (Junior and Senior High School Forms) prepared by Dr. Ross Mooney of the Ohio State University.

TABLE 2

CHANGES IN SPECIFIC RELIGIOUS BELIEFS DURING ADOLESCENCE AS SHOWN BY THE PERCENTAGE OF 12, 15, AND 18 YEAR-OLD CHILDREN WHO CHECKED VARIOUS STATEMENTS INDICATING (*a*) BELIEF, (*b*) DISBELIEF, OR (*c*) UNCERTAINTY (WONDER)*

Statement	"Believe"			"Not Believe"			"Wonder About"		
	12	*15*	*18*	*12*	*15*	*18*	*12*	*15*	*18*
God is a strange power working for good, rather than a person	46	49	57	31	33	21	20	14	15
God is someone who watches you to see that you behave yourself, and who punishes you if you are not good	70	49	33	18	37	48	11	13	18
I know there is a God	94	80	79	3	5	2	2	14	16
Catholics, Jews and Protestants are equally good	67	79	86	9	9	7	24	11	7
There is a heaven	82	78	74	4	5	5	13	16	20
Only good people go to heaven	72	45	33	15	27	32	13	27	34
Hell is a place where you are punished for your sins on earth	70	49	35	16	21	30	13	27	34
Heaven is here on earth	12	13	14	69	57	52	18	28	32
People who go to church are better than people who do not go to church	46	26	15	37	53	74	17	21	11
Young people should belong to the same church as their parents	77	56	43	13	33	46	10	11	11
The main reason for going to church is to worship God	88	80	79	6	12	15	4	7	6
It is not necessary to attend church to be a Christian	42	62	67	38	23	24	18	15	8
Only our soul lives after death	72	63	61	9	11	6	18	25	31
Good people say prayers regularly	78	57	47	9	29	26	13	13	27
Prayers are answered	76	69	65	3	5	8	21	25	27
Prayers are a source of help in times of trouble.	74	80	83	11	8	7	15	10	9
Prayers are to make up for something that you have done that is wrong	47	24	21	35	58	69	18	17	9
Every word in the Bible is true	79	51	34	6	16	23	15	31	43
It is sinful to doubt the Bible	62	42	27	18	31	44	20	26	28

*Discrepancies between the totals of "Believe," "Not Believe," and "Wonder About," and 100 per cent represent the percentages who did not respond to the statements. Differences of 8 or 9 will ordinarily yield a *CR* of 2.0, depending upon the magnitude of the percentages involved

ber of rather significant changes have occurred. First to be noted is the fairly marked discarding of such beliefs as "Every word in the Bible is true," "It is sinful to doubt the Bible," "God is someone who watches you to see that you behave yourself, and who punishes you if you are not good." Roughly two-thirds (close to 70 per cent) of the twelve-year olds believed these statements to be true, only a third or less of the eighteen-year olds would agree. Other statements which were notably less frequently believed by 18 than 12 year olds are "Only good people go to heaven," "Hell is a place where you are punished for your sins on earth," and "Prayers are to make up for something wrong you have done" Thus many rather specific beliefs taught to, or picked up by, young children are no longer held by most of those in the late teens

Second, a greater tolerance with respect to religious beliefs and practices is apparent with development into adolescence. Thus *more* eighteen than twelve-year olds agree that "Catholics, Jews, and Protestants are equally good," that "It is not necessary to attend church to be a Christian" Fewer eighteen than twelve-year olds believe that "People who go to church are better than people who do not go to church," that "Young people should belong to the same church as their parents," and that "Good people say prayers regularly"

Third, it will be noted that drops in the proportion "believing" a statement are not necessarily compensated for by increases in the proportion "not believing" that statement. With respect to certain beliefs, at least, the issues are far from settled even by 18 years of age More of the eighteen-year old group, for example, "wonder about" the statement "Every word in the Bible is true," than who agree with the statement or who disbelieve it. In the case of 12 of the 19 statements contained in Table 2, there is a trend toward increased "wondering." Such is true of beliefs involving the hereafter (death, heaven, and hell), certain concepts of God, certain beliefs with respect to prayer and belief in the Bible.

It is evident from these data that the religious views of adolescents in the late teens differ from those of twelve-year old children in a number of very significant ways. In fact, of the 52 statements included in this study, changes were great enough in 36 statements to be two or more times their standard errors[8] That even the eighteen-year old group have not yet reached a satisfactory religious philosophy is suggested by the amount of uncertainty that exists with respect to various beliefs and concepts. However, it must be admitted that many a person with a "satisfactory" set of religious beliefs would check "wonder about" when asked to react in one of three stated ways to statements such as those of this study

RELIGIOUS DOUBTS AND PROBLEMS

It might be expected in view of the data thus far presented that there would be a general increase during development in adolescence in the proportion of statements "wondered about" and probably also in the number of problems faced. Both expectations are in line with the beliefs apparently held by many writers in the field of adolescent psychology. Two lines of evidences are presented regarding this question. First a "wonder score" was obtained for each subject by counting the number of items he checked as "wondering about." Second, a "problem score" was obtained by an analysis of the responses of each subject to the second portion of the questionnaire. Responses to the problems there listed were summarized by giving weights of zero, one, and two respectively to responses of "never," "sometimes," and "often," and summing the 18 responses. The results of such analyses are summarized in Table 3.

---

[8] It would have been desirable to study age changes for particular groups, especially Catholic and Non-Catholic, but there were too few cases in any particular category to make such analysis significant.

TABLE 3

TRENDS DURING ADOLESCENCE IN THE GENERAL EXTENT OF DOUBT AND CONCERN REGARDING RELIGIOUS ISSUES AS SHOWN BY MEAN "WONDER" AND "PROBLEM" SCORES OF THE THREE AGE GROUPS

	Age			CR_{Diff}
	12	15	18	12-18
Wonder scores	7.9	8.1	8.7	1.1
Problem scores	9.7	9.9	9.9	0.3
No. of cases	174	243	130	

A slight trend is noticeable in the case of the average number of statements wondered about, but the difference between the twelve-year old and the eighteen-year old groups is so small compared to the amount of sampling error probably present that it cannot safely be considered more than a chance difference. In the case of the "problem scores" no difference appeared in the means for the three age groups. Thus insofar as these findings are concerned the hypothesis that adolescence is a period of *generally* increased religious doubts and problems is clearly not substantiated. Since age trends here (and also sex differences) are slight, it was decided to compare Catholics with non-Catholics by grouping sixth, ninth, and twelfth grades and both sexes together. When this was done, it became apparent that Catholics had lower scores in both "wonders" and "problems" than did the non-Catholics. In the case of "wonder scores" the mean difference was 3.0 (*CR* 4.9); in the case of "problem scores" the mean difference was 1.2 (*CR* 2.2).

That age trends exist with respect to particular problem faced is to be expected, and appropriate data revealing such trends are presented in Table 4. Over half of the eighteen-year old group of subjects indicate that the following problems trouble them sometimes or often: Disliking church service,

TABLE 4

FREQUENCY WITH WHICH PARTICULAR RELIGIOUS PROBLEMS EXIST AT VARIOUS AGES THROUGH ADOLESCENCE AS SHOWN BY PERCENTAGE OF DIFFERENT AGE GROUPS WHO CHECK EACH PROBLEM AS SOMETIMES OR OFTEN PRESENT

Problem	Age 12	 15	 18	CR_{Diff} 12-18
Having a different religion from other people	34	25	27	1.3
Disliking church service	33	47	60	4.9
Being forced to go to church	30	31	27	0.6
Disliking parents' religion	11	8	12	0.7
Failing to go to church	67	67	67	0.0
Changing my idea of God	29	25	31	0.4
Losing faith in religion	27	32	31	0.8
Doubting prayer will bring good	37	44	35	0.4
Getting help on religious problems	53	54	56	0.6
Choosing a religion	21	20	15	1.5
Parents objection to church membership	23	14	11	2.6
Wanting to know the meaning of religion	53	48	60	1.2
Wanting communion with God	59	47	57	0.3
Heaven and hell	53	53	66	2.3
Sin	71	62	72	0.2
Conflicts of science and religion	42	50	57	2.6
Being teased about my religious feelings	26	22	18	1.5
Wondering what becomes of people when they died	67	56	80	2.3
Number of cases	174	243	130	

Failing to go to church, Getting help on religious problems, Wanting communion with God, Wanting to know the meaning of religion, Heaven and Hell, Sin, Conflicts of science and religion, Wondering what becomes of people when they die. It would seem as though adolescents do have certain religious problems regarding which they would like help, but apparently find conventional religious programs unsatisfying. Three of the problems listed—Disliking church services, Heaven and Hell, and Wondering what becomes of people when they die—become more prevalent with increased age, the critical ratios of the twelve to eighteen-year old differences in these cases being over 2.0. One problem—Parents objection to church membership—decreases sufficiently to show fair statistical reliability. Just what is meant by this response is not clear. It may be that parents of twelve-year olds consider them too young to join a church.

DISCUSSION

It is not implied by the selection of age groups of varying statuses with respect to pubescence, that the differences shown are a function of pubescence. Rather, it would seem more reasonable to assume that they are the result of accumulated experience in combination with increasing intellectual maturity which makes the adolescent more capable of interpreting the environment of ideas and facts in which he is becoming increasingly immersed. Greater intellectual maturity might be expected to increase sensitivity to inconsistencies either among the beliefs and views an individual contacts, or between his already established beliefs and new learnings Also with greater maturity the adolescent is more capable of abstract generalizations which might result in discarding some specific beliefs in favor of more general ones.

Do such findings have any practical value? To the psychologist and others interested in human development such data are of interest as descriptions of how development proceeds But are there any implications for those interested in religious education? Two obvious implications may be mentioned First, those issues represented by statements which are increasingly "wondered about" as age increases may give clues as to appropriate topics for consideration in the teen years in both Sunday School classes and young people's groups Second, beliefs discarded by children as they grow older may well be studied for their implications for teaching at earlier ages. Children's concepts regarding religion are more concrete and specific than are those of adults, the latter tending to be abstract and general. This change represents the normal growth of concepts. It would seem desirable that the specific and concrete beliefs taught to children be beliefs compatible with the more abstract adult views, and not beliefs later to be discarded because of incompatibility.

A few comments now regarding methods of studying adolescence. In studies of adolescent development it is necessary to include a sufficient age range in the sample to give perspective to the findings. Because a group of high school students have certain characteristics is no indication that *those characteristics are peculiar to adolescence* It may well be that young children and adults as well show the same characteristics. The need for such an age perspective is interestingly demonstrated in data collected by P. M. Symonds and published by Shuttleworth (8, Fig 317). "Philosophy of life" which is ordinarily assumed to be a problem and interest typical of adolescents, was demonstrated to be of more interest *at 50* than at 15. The present study has not involved such a wide age range as did Symonds', but some perspective is given. The situation in adult years needs also to be studied.

However, certain difficulties are encountered in attempting to study wide age ranges by comparable procedures It is desirable to get reactions of different age groups to essentially the *same situation*. But ten, eleven, or twelve-year old youngsters may be unable to comprehend many statements entirely clear to adolescents in the middle or late teens. The question may well be raised regarding the present study as to whether the youngest group could adequately comprehend the questionnaire used. The writers believe that their comprehension was adequate. No more of the youngest group than the middle or oldest group failed to respond to the statements, and the responses of the twelve-year olds were, on the average, either comparable to the older age groups or continued trends apparent in the fifteen and eighteen-year old comparisons.

It is worthwhile also to point out that in research of this sort, average scores may very well obscure more facts than they reveal This is the writer's criticism of Dimock's study. In contrast to his conclusion of the relatively static nature of religious ideas in adolescence, the present findings reveal rather significant shifts during the ages considered The need for detailed analysis is revealed even within the present study. Average "wonder scores" and "problem scores" showed little change. Yet analysis of responses to specific beliefs and problems indicated changes of high statistical reliability. The meaning of any "score" rests to a considerable extent upon the particular questions or items which that score summarizes. When items showing increasing frequency of checking are added in with other items of a decreasing trend, the differences cancel and a fairly stable over-all score may conceal the actual change

It is probable that the use of a questionnaire dealing exclusively with religious issues has momentarily focused the subjects' attention upon a rather narrow range of human experience, and by taking these experiences out of the context of everyday living may give a biased picture of the importance of such issues and problem in the lives of adolescents. Three of the problem items used in the present study were included in a modified form of the Mooney *Problem Check List* (Junior High Level) used by the senior writer in another study Although 22, 48, and 25 per cent of ninth graders in the present study checked these items as bothering them sometimes or often, only 9, 18, and 20 per cent of a group of 100 ninth graders so checked the same items when they were included among a great variety of other kinds of problems. This is simply another illustration of the part the instrument used plays in determining the "facts" discovered, and warns against too literal an interpretation of the absolute figures presented in this paper The differences noted and trends revealed, however, may more safely be depended upon[1]

SUMMARY

Five hundred forty-seven children and adolescents, in three groups which averaged 12, 15, and 18 years of age, responded to a questionnaire which listed 52 statements representing various religious beliefs and 18 problems dealing with religious issues Many significant differences appeared in religious beliefs when twelve-year olds and eighteen-year olds were compared A greater tolerance with respect to religious beliefs and practice, a discarding of a number of specific beliefs and increased "wondering about" statements regarding the hereafter (death, heaven, hell) constituted the major trends An analysis of responses of "wondering about" particular beliefs and "problems" did not substantiate the commonly accepted hypothesis that adolescence is a period of generally increased religious doubts and problems Catholics "wondered about" fewer beliefs and checked fewer problems than did non-Catholics However, the specific problems checked by the subjects indicated that many do have problems of a religious sort, want help (in fact, one of the greatest problems seemed to be getting such help), but are dissatisfied with conventional church services

REFERENCES

1 DIMOCK, H. S New light on adolescent religion *Relig. Educ*, 1936, **31**, 273-279

2 ———— Rediscovering the Adolescent New York· Association Press, 1937.

3. DUDYCHA, G. J Religious beliefs of college students *J Appl Psychol*, 1933, **17**, 585-603.

4 FRANZBLAU, A N Religious beliefs and character among Jewish adolescents. *Teach Coll., Columbia Univ, Contrib Educ*, 1934, No 634

5 KATZ, D, & ALLPORT, F H Students Attitudes. Syracuse Craftsman Press, 1931

6 MACLEAN, A N Idea of God in Protestant religious education *Teach Coll, Columbia Univ, Contrib Educ*, 1930, No 410

7. MOONEY, R Surveying high school students problems by means of a problem check list *Educ Res Bull.*, 1942, **21**, 57-69

---

[1] One report by Mooney (7) suggests that religious and moral problems are relatively minor in importance in adolescence when compared with other types of problems

8 Shuttleworth, F. R The adolescent period: a graphic and pictorial atlas *Monog Soc Res Child Devel*, 1938, 3, No 3 Pp 246

School of Education
Syracuse University
Syracuse, New York

Goucher College
Baltimore, Maryland

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1944, 65, 301-303

AN APPARENT CASE OF NEUROTIC INFLUENCE ON SOMATIC FUNCTION*

Laboratory for Child Research, Mooseheart, Illinois

MARTIN L. REYMERT

This case is presented to add to the case literature accumulating on this topic. The subject was the eight-year-old son of personal friends of the writer. The father is a lawyer with much dramatic experience during his college career. The mother (who made the material available, including the radiologist's X-rays and the pediatrician's report, is a trained speech pathologist.

In 1936, George Bernard Shaw's *Candida* was being produced by local theatrical talent under the direction of a Mr. P., a professional from a neighboring town. The mother of the boy played the title rôle while his father played the part of her husband, Morell, a minister. The director, Mr P , took the part of Marchbanks, the poet in the play

The subject, Bobby, was taken with his parents to rehearsals of the play on several Sunday afternoons He knew the members of the cast by their character-names and according to the mother "adored Mr. P. (Marchbanks) whom he had also met more intimately in his parental home when an occasional Sunday night supper was served."

In the late afternoon, a few hours before the performance was to be given, Bobby became very ill, vomiting violently. Before the family had to leave for the auditorium, the mother called the pediatrician in whose care Bobby had been since birth. He suggested that the boy be brought to him for examination if the condition did not show improvement by morning. The boy pleaded with the mother not to go to the play. However, the play did go on. Both parents were anxious about the child as by evening he

---

*Received in the Editorial Office on February 2, 1943

was suffering more and was in great pain. Backstage telephone calls kept them in touch with the home.

The boy was asleep when the parents reached home that night. The vomiting continued in the morning when he tried to eat anything so he was taken to the doctor who made an examination of the child. The blood count was 10,900 white cells, 67 polynuclear, no segmented cells, and 2 eosinophiles An X-ray was taken that morning of the gall bladder Another one was taken later in the day after the mother had had the child home and fed him rich foods such as eggs, malted milk, etc The report from the radiologist, dated November 20, 1936, is as follows· Film of gall-bladder after oral dye shows elongated organ densely filled with dye (A). Film after fatty meal shows it greatly reduced in size (B). The conclusion of the report was "Enlarged gall-bladder with very active function." After about four days the digestive system began to function normally again.

The cause of the disturbance was unknown to the mother although she conjectured that possibly some terrific emotional disturbance had been responsible for the physical upset. Since it occurred on the afternoon of the play's presentation, the play might offer some clue. Nothing was said to the child but subsequently press clippings were read aloud and *Candida* was discussed when friends called Thus the topic was kept fresh and alive. Several days after, Bobby said to his mother. "Mother, I like Marchbanks I think he's a swell guy and I don't want you to say anything to him, but I'm glad you didn't choose him instead of my Daddy—and before all of those people!" This statement by the subject probably gives the key to his difficulty. In the third act of the play there is a scene where Candida "chooses" between Morell and Marchbanks. It is quite a dramatic and tense moment in the play. Now the child had seen the rehearsal and knew that on every such occasion the mother (Candida) had always chosen Morell (the subject's father) However, anxiety over the possibility that she might choose otherwise on the night of the play, "and before all of those people" probably disturbed the child sufficiently to give rise to the somatic symptoms described here His own liking for Marchbanks made him feel all the more keenly and realistically the difficulty of the "choice" his mother, as Candida, had to make and this factor probably increased his anxiety. His pleading with his mother not to go to the performance is consistent with this hypothesis. The fact that his difficulty was climaxed at about the time of the performance and that subsequently, when he learned everything had gone well, the difficulty cleared up, further strengthens this interpretation

Thus, while the evidence cited above is not conclusive proof that we have here a case of "neurotic influence on somatic function," all facts seem to fit such a conclusion

Laboratory for Child Research
Mooseheart, Illinois

The Journal of Genetic Psychology, 1944, 65, 305-316

BOOKS

The *Journal of Genetic Psychology,* the *Journal of General Psychology,* and the *Journal of Social Psychology,* will buy competent reviews at not less than $2 per printed page and not more than $3 per printed page, but not more than $15 00 for a single review.

Conditions. Only those books that are listed below in this section are eligible for such reviews In general, any book so listed contains one or more of the following traits. (*a*) Makes an important theoretical contribution; (*b*) consists largely of original experimental research, (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology, (*d*) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph D degree or its equal in training and experience

Procedure If among the books listed below there is one that seems important to you, you are invited to write a review of that book It is not necessary to make arrangements with the Editor. *Just send in your review It does not matter if* the book in question has been reviewed before

(1942)

SHELDON, W. H. The Varieties of Temperament New York Harper, 1942 Pp 520

KLOPFER, B, & KELLEY, D. M. The Rorschach Technique. New York· World Book, 1942 Pp. 436

SCHILDER, P. Goals and Desires of Man New York Columbia Univ Press, 1942 Pp. 305

MILLER, J G Unconsciousness New York Wiley, 1942 Pp. 329

ALLPORT, G. W *The Use of Personal Documents in Psychological Science* New York. Social Science Research Council, 1942 Pp 210.

BENDER, I E, IMUS, H A, ROTHNEY, J. W M, KEMPLE, C., & ENGLAND, M R Motivation and Visual Factors Hanover, N. H · Dartmouth College, 1942 Pp 369

BORING, E. G Sensation and Perception in the History of Experimental Psychology New York Appleton-Century, 1942 Pp 644

GROVES, E R, GROVES, G. H, & GROVES, C Sex Fulfillment in Marriage New York: Emerson, 1942. Pp 319

WATSON, G (*Ed*) Civilian Morale New York Houghton Mifflin, 1942 Pp 463.

BIBER, B., MURPHY, L B, WOODCOCK, L. P, & BLACK, I S Child Life in School New York Dutton, 1942. Pp 644

HOLLINGWORTH, L S Children Above 180 IQ New York World Book, 1942 Pp. 332

KLUVER, H (*Ed*) Biological Symposia Visual Mechanisms Lancaster. Cattell Press, 1942. Pp 322

LANDIS, C, & BOLLES, M. M Personality and Sexuality of the Physically Handicapped Woman New York Hoeber, 1942. Pp 171

LEWIS, T Pain New York· Macmillan, 1942 Pp. 192

McGeoch, J A. The Psychology of Human Learning. New York: Longmans, Green, 1942. Pp 633

Schilder, P. Mind. Perception and Thought in Their Constructive Aspects New York. Columbia Univ. Press, 1942. Pp. 432

Rees, H. E. A Psychology of Artistic Creation. New York: Teach. Coll., Columbia Univ, 1942. Pp. 209.

Seashore, C E. Pioneering in Psychology Iowa City: Univ. Iowa Press, 1942

Rapaport, D. Emotions and Memory Baltimore: Williams & Wilkins, 1942 Pp. 282.

(1943)

Gesell, A, & Ilg, F. L. Infant and Child in the Culture of Today New York Harper, 1943. Pp. 399.

Leeper, R. W Lewin's Topological and Vector Psychology. Eugene: Univ. Oregon 1943 Pp 218

Yerkes, R M Chimpanzees. A Laboratory Colony. New Haven Yale Univ Press, 1943 Pp 321.

Young, P T. Emotion in Man and Animal New York: Wiley, 1943 Pp 422

Masserman, J. H. Behavior and Neurosis. Chicago Univ Chicago Press, 1943. Pp 269

May, M. A A Social Psychology of War and Peace New Haven. Yale Univ. Press, 1943 Pp 281

Wolff, W The Expression of Personality. New York: Harper, 1943 Pp. 334

Estabrooks, G. H Hypnotism. New York. Dutton, 1943. Pp. 249.

Barker, R. G., Kounin, J S, & Wright, H. F Child Behavior and Development. New York McGraw-Hill, 1943 Pp. 652

Newcomb, T M Personality and Social Change. New York: Dryden, 1943 Pp 225

McGraw, M B. The Neuromuscular Maturation of the Human Infant New York. Columbia Univ Press, 1943. Pp 140.

Child, I L Italian or American? The Second Generation in Conflict. New Haven. Yale Univ. Press, 1943 Pp. 208.

Griffith, C R. Principles of Systematic Psychology. Urbana, Ill: Univ Illinois Press, 1943. Pp. 718.

Meier, N. C Military Psychology. New York: Harper, 1943. Pp. 395

Strong, E K., Jr. Vocational Interests of Men and Women Stanford University Stanford Univ Press, 1943. Pp. 746.

Tomkins, S S. (*Ed*) Contemporary Psychopathology. Cambridge: Harvard Univ. Press, 1943. Pp. 600

Hull, C. L Principles of Behavior. New York Appleton-Century, 1943. Pp 422

(1944)

Cantril, H Gauging Public Opinion Princeton: Princeton Univ. Press, 1944 Pp 318

Hunt, J M (*Ed.*) Personality and the Behavior Disorders (2 vols) New York Ronald Press, 1944. Pp 1242.

CRITICAL REVIEWS OF RECENT BOOKS

The Journal of Genetic Psychology, 1944, **65**, 307-313.

(*Gesell, A., & Ilg, F L, in collaboration with Learned, J., & Ames, L B. Infant and Child in the Culture of Today. New York Harper, 1943 Pp 399*)

REVIEWED BY DOROTHY TILDEN SPOERL

Infant and Child in the Culture of Today, in spite of certain theoretical weaknesses and omissions to be discussed later, fulfills a long felt want in the field of books for parents and other laymen interested in the problems of child psychology. Written largely in non-technical terms it is at the same time addressed to thinking persons and not "written-down" to a normative level as are so many of the books available for parent education.

For the psychologist who was asked by a parent, "what shall I read that will help me with my child," for the person who is interested in working with children and wants to know, "what will help me differentiate the interests and activities at various levels," for the medical person (including the dentist, the person working with vision as well as the pediatrician) who is interested in "what are the psychological factors accompanying physical maturation" this is a book which can be recommended. And although one may feel, with the present reviewer, that it leaves much to be desired, one can say at the same time that among books of this kind it is one of the best

The volume itself is divided into three main sections the first is concerned with *Growth and Culture*, the second with *The Growing Child*, the last with *The Guidance of Growth*. Although somewhat inter-related one somehow gets the impression that perhaps the sections were not worked on in all cases by the same members of the group of four who are responsible for the book as a whole Because of differences in approach and materials the three sections will be considered somewhat separately in review.

GROWTH AND CULTURE

This section seems to be far the weakest in the book It deals in rather generalized terms with various problems of growth, personality, individuality, and culture, and often falls back upon a mere cataloguing of stages rather than an understanding explanation of their essential nature. If one were to state the most apparent weakness of approach (both in this section and in the others) it would have to be the almost complete lack of any adequate theory of motivation. Indeed one might quite accurately state that the authors seem to have no theory of motivation at all.

In considering the manner in which the problems of personality and individuality are presented one is struck by the strong behavioristic attitude which pervades the writing Emotions and emotional experience seem to be outside of the realm of matters of real importance. For instance, when in a discussion on the growth of the mind (p 20) the authors consider the three levels of reality which mental life embraces they are the levels of the vegetative functions, the world of things, and the world of persons Certainly beyond this there are at least the world of ideas, and the world of the spirit, and to the clinically oriented person the world, which if not separate, at least colors all of the others, the world of emotions.

Perhaps this ignoring of emotions is best illustrated in the discussion of personality, where are found these two statements:

> These mechanisms (those of developmental physiology) are so lawful and so fundamental that children of similar chronological age are in general most comparable with respect to their emotional characteristics (p. 37).
>
> We wish to avoid the suggestion that the personality of infant and child is a product of emotions which operate in some mysterious way through the subconscious or otherwise We should prefer to think of personality in less mysterious and more realistic terms as a structured end product of the child's developmental past (p 37).

Surely one cannot, if conversant with the literature of psychoanalysis or with clinical records of abnormal and deviant behavior, believe but that emotional characteristics are one of the commonest, even the basic, cause of deviating behavior. It is, after all, the emotional complexes of the child which cause uneven development, abnormal behavior, and the problems which arise in the course of child maturation; and surely from this point of view emotional development since it is so important in the determining of behavior (at least outside of Yale) must be the field in which one finds the least rather than the most comparability so far as children are concerned at any given age level.

In this connection, and particularly with the second quotation in mind, one must raise the question as to how the child's developmental past can ever be considered apart from his emotional past? How can one ignore in the structuring of development the basic foundation of the structure, which must always be the strength and directions of the child's emotional life? Whatever one may think of the details of Freudian psychology (which incidentally are ignored completely in this book, the one mention of Freud being a reference to his "suggestion that dreams are the guardians of sleep"), it does seem that he, together with his followers and those who have used and modified his theories, has established the fact that emotions, and particularly the emotions of childhood, have a permanent and a fundamental effect upon the personality of any individual, whether he be infant, child, or adult.

Indeed one often wonders while reading *Infant and Child in the Culture of Today* to what extent the authors feel that the physical life of the child *is* the personality of the child The statement is made in referring to the Behavior Profiles to be presented later in the book that, "this profile sketches the personality characteristics of a typical child of a given maturity" (p. 38), yet the characteristics most completely discussed are physical characteristics, and certain personal-social characteristics. The preponderance of space, however, goes to the physical characteristics. And the personal-social characteristics seem to be more those having to do with the process of acculturation than that of personality development Unless, of course, personality is primarily physical.

One other feature of this section of the book should be mentioned. If the laymen can grasp the full import of the theory of "self-demand schedules" they are an excellent device Suggesting that following the clock is a dangerous thing; and that the parent who will follow the child and see the individual variations can work out a pattern for the behavior day of that particular child, the authors show how in this way the parent will become more understanding of the nature of the child and the schedule which is optimal for him. Fluctuations in behavior are described as often being "forward thrusts" which the child is attempting The danger, for the laymen, is in accepting the fluctuations as the demands and perpetuating them, but it is a danger which can be somewhat alleviated by a careful study of section two.

The Growing Child

This section of *Infant and Child in the Culture of Today* through a clear cut and ingeneous system, presents the factors of the physical and mental (but not the emotional) development of the child. It is the section of the book which makes one feel that the book as a whole is desirable for recommendation to parents and others in search of material that will enhance their understanding of childhood.

Up to 18 months of age the material is presented through the "Behavior Profile" and the "Behavior Day." After that level there is added "Cultural and Creative Activities," "Nursery Behavior," and "Nursery Techniques." The Behavior Profile is a narrative description of the general level of the child designed to "give us a picture of the kind of child with which the culture has to deal at a certain stage of his maturity." There seems to be little in the profiles which is not included in a clearer and more definitive form in the descriptions of the Behavior Day It does, however, have a basic purpose of orientation toward the given level which is sufficiently fulfilled to be helpful to those not conversant with the developmental aspects of child life.

The Behavior Day is a typical outline of the manner in which a "somewhat typical child" advances. Starting with a description of the physical events of the day from the time of waking, through naps, play, excursions, meals, routines, and back to sleep, it gives an excellent idea of what may be expected at various age levels. The reader needs always to bear in mind that, as the authors so often state, these are typical days, they are not meant to be set up as a guide for the behavior day of a specific infant or child. Following this outline there is a section in which various specific factors are considered: Sleep, Feeding, Elimination, Bath and Dressing, Self-Activity, and Sociality. Again it is noticeable that the emotional aspects are largely missing. Indeed one feels that the orientation is much more physical than psychological, as though the medically trained authors have often superseded the psychologically trained But withal it is a very real help to the parent in search of specific suggestion for general guidance.

One thing which should prove especially helpful is the many explanations as to the cause of what would otherwise seem to be deviant behavior. To select one or two at random should be sufficient for illustration. It is pointed out that at 21 months there are frequently relapses in bowel training for the previously well trained infant. The explanation is offered that this is related to a new "but too powerful release mechanism" If the parent who is upset about this apparently wilful regression can be shown the cause and the normality of the incident he is able to accept what would otherwise become a focal point for emotional turmoil. Under social development at 30 months is considered the "ritualism" which is said to be characteristic of this age Again it should help parents to accept the child's dislike of change if he can understand the mechanism which underlies it Imaginative devices, such as imaginary companions and shifts in identity, are considered under the three-year-old level The treatment is somewhat superficial, being more or less dismissed with the statement "probably the intricacies of individual emotional development are being worked out by the child through these imaginative devices." While for a real understanding of these phenomena it is necessary to consider carefully the deep-lying dynamic processes which mediate emotional development, it will, nevertheless, be helpful for the average parent if he can realize the frequency with which these devices develop in the life of children But, in a book of this type which is so apt to become "THE BOOK" by which a child is brought up there should have been a more adequate presentation of the emotional factors which cause the regression, implement the ritualism, give birth to the imaginary companions, and can accrue to bowel control Again, the omission of an adequate theory of emotions and of motivation has seriously handicapped what would otherwise have been invaluable material. It retains still a great deal of usefulness.

The other materials presented need little discussion "Cultural and Creative Activities" are for the most part a listing of the manner in which a child will demonstrate his interest in books, music, painting, etc They are suggestive, and should prove helpful to the parent who is wondering "what the child is ready for now" The section on "Nursery Behavior" shows how the behavior of the behavior day will be modified by the transition from home to school. The section on "Nursery Techniques," while it will hardly be adequate for those who are actually dealing with children in a nursery school situation, will be helpful to parents who may wish to use the techniques at home

THE GUIDANCE OF GROWTH

To the mind of the reviewer this section, which by and large merely reiterates in briefer form many of the findings already discussed, adds little to the value of the book. And in its over-simplifications it constitutes, considered as a whole, one of the best examples of the lack of consideration of the problems of emotion and motivation already discussed.

The growth complex is held to have three outstanding features: the day cycle, the self-regulatory fluctuations, and constitutional individuality. If the latter were not dismissed in a partial page it might have contained some of the materials which the book so sadly lacks.

The specific discussion of the growth complex is divided into sections concerned with what seem, apparently, to the authors to be its important divisions: sleep, feeding, bowel control, bladder control, personal and sex interests, and self-activity, sociality, self-containment. In the discussion of all of these the general normal development of each is charted, the total impression being that the sequential patterning is an extremely orderly process. In most cases a brief consideration of deviant behavior is included. Unfortunately for the most part this deviant behavior is dismissed as being a deviation from the normal developmental sequence, which if ignored will again reach a stage of equilibrium

For example these two brief quotations from the section on personal and sex interests

> Indeed, so integrated and unified are the processes of psychological growth that an intelligent management of feeding, sleep, and elimination behavior, in itself, almost insures a favorable organization of other fields of behavior The problems of so-called sex-hygiene may be approached with confidence, and without an undue sense of mystery (p 334)

There are many stepping stones along the way,—articulate and inarticulate Here are some of them:

> Johnny,—that's me. * I am I That is my mother. * That is my father * He is a man. * I am a boy * Susan is a girl * She has a father and mother too * I was a baby. * I grew. * I came from my mother I am going to get bigger. * I am going to go to school * (p 334)

Then later these thirteen statements are referred to again:

> And so the spiral once more comes to a full turn By slow and not altogether painless stages the 5-year-old has achieved the thirteen sequential judgments which register his cultural maturity (pp 340-341)

Certainly anyone would agree that the problems of sex-hygiene can, and should be, approached with "confidence, and without an undue sense of mystery" But they do not come alone through these "thirteen sequential judgments," nor, in many cases without emotional crises which end in motivational deviations that are of lasting effect If the child were brought up in his home alone, and that home were near to perfection, and the dictates of *Infant and Child* were followed with care, this might be true But the child has friends, a community, and experiences that we might wish to save him from. He hears stories from his classmates, he reads pornographic literature, he gets a sense of guilt from his sexual experimentations, his social efforts are not received, and the 13 stages do not follow in orderly fashion. Or, if they did, he is emotionally unable to accept them as they stand. And the deviant behavior that results is not a simple deviation from the growth *complex, it is something which thwarts and twists that growth* complex *in an* enduring manner

Referring to deviations in bowel control, to such problems as thumb sucking, etc, always the answer seems to be that they will run a general sequential patterning, and that when the pattern has been completed they will tend *to disappear Almost never is any reference made to the emotional and* motivational factors which cause them in the first place, continue them in the second, and more often than not in the last analyses cause them not to disappear but rather to be transformed into some other and more serious form of deviation

It is for these reasons that the present reviewer feels that it is unfortunate that section two of the book does not stand alone, for it is this section which will give it its lasting values. Many parents who try to assimilate the whole volume, and who accept too literally the doctrines of this last section will be led into serious difficulties in the rearing of their children. It is not that the book contains *statements that are not true, it is rather that the over*-simplifications and more importantly the omissions, tend to lead to misconceptions on the part of the uninitiated layman.

American International College
Springfield, Massachusetts

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i e., the most recently received books will be found at the end of the list.)

Graubard, M Man's Food. Its Rhyme or Reason New York: Macmillan, 1943 Pp 213.

Williams, C. T. These We Teach. Minneapolis Univ Minnesota Press, 1943 Pp 188

Rivlin, H. N, & Schueler, H. (*Eds*) Encyclopedia of Modern Education New York: Philosophical Library, 1943. Pp 902

Clayton, A S. Emergent Mind and Education New York: Columbia University, 1943 Pp 179.

Srinivas, M N. Marriage and Family in Mysore. Bombay New Book, 1942. Pp 219

Meier, N C Military Psychology New York Harper, 1943 Pp 395

Baxter, B, & Cassidy, R Group Experience. New York Harper, 1943. Pp. 218

Whiteway, H. L. Scientific Method and the Conditions of Social Intelligence New York Columbia University, 1943. Pp 188

Jessup, B E. Relational Value Meanings. Eugene, Oregon: Univ Oregon Press, 1943. Pp 175.

Glass, B Genes and the Man New York Columbia University, 1943 Pp 386

Jennings, H H Leadership and Isolation. New York. Longmans, 1943. Pp. 240

Lee, A M, & Humphrey, N D. Race Riot. New York Dryden, 1943 Pp. 143.

Hollingworth, H. L Leta Stetter Hollingworth Lincoln, Neb · Univ Nebraska Press, 1943. Pp 204

Lowinger, A The Methodology of Pierre Duhem New York: Columbia Univ Press, 1941. Pp 184

Taliaferro, W. H, *Ed* Medicine and the War. Chicago· Univ. Chicago Press, 1944 Pp 193

Hull, C L. Principles of Behavior New York: Appleton-Century, 1943 Pp 422

Nash, A S The University and the Modern World. New York: Macmillan, 1944 Pp. 312.

Kasanin, J. S Language and Thought in Schizophrenia Berkeley· Univ. California Press, 1944. Pp 133.

Bowley, A H. Guiding the Normal Child New York· Philosophical Library, 1943 Pp. 174.

Cantril, H Gauging Public Opinion Princeton. Princeton Univ Press, 1944. Pp, 318.

Flesch, R Marks of Readable Style. New York: Columbia University, 1943. Pp 69

Sargent, S S The Basic Teachings of the Great Psychologists New York: New Home Library, 1944. Pp 346.

Hunt, J M, *Ed* Personality and the Behavior Disorders (2 vols) New York: Ronald Press, 1944 Pp. 1242

Klein, D. B Mental Hygiene. New York Holt, 1944. Pp 498.

Salter, A. What is Hypnotism? New York: Smith, 1944. Pp 88

Barnouw, A. J, & Landheer, B The Contribution of Holland to the Sciences New York· Querido, 1943 Pp. 373

Moore, U., & Callahan, C. C Law and Learning Theory. New Haven Yale Law Journal, 1943 Pp. 136

Ribble, M A The Rights of Infants. New York: Columbia Univ. Press, 1943. Pp 118

Parker, C M. Steel in Action Lancaster, Pa: Cattell Press, 1943 Pp 221

Woodworth, R S, & Sheehan, M. R First Course in Psychology New York Holt, 1944. Pp 445

Thorndike, E L Man and His Works Cambridge. Harvard Univ Press, 1943. Pp 212

Winn, R B, *Ed.* Encyclopedia of Child Guidance. New York: Philosophical Library, 1943 Pp 456

Boyd, W. Evacuation in Scotland London: Univ London Press, 1944 Pp. 224

Cole, L Attaining Maturity. New York Farrar & Rinehart, 1944. Pp. 212.

Lindner, R. M Rebel without a Cause. New York: Grune & Stratton, 1944. Pp 296

Allport, G. W The ABC's of Scapegoating Chicago Central Y M C A College, 1944 Pp 72

Klein, D. B. Mental Hygiene New York: Holt, 1944 Pp. 498.

Benne, K D A Conception of Authority New York: Columbia Univ., 1943 Pp 227

Snyder, H E Educational Inbreeding New York: Columbia Univ., 1943. Pp 160

American Psychiatric Association One Hundred Years of American Psychiatry New York: Columbia Univ Press, 1944 Pp 649

Du Bois, C The People of Alor Minneapolis: Univ. Minnesota Press, 1944 Pp. 654.

Pigors, P, & Pigors, F. Human Aspects of Multiple Shift Operations Cambridge Mass Inst Tech, 1944 Pp. 90

Gallup, G. A Guide to Public Opinion Polls. Princeton Princeton Univ Press, 1944. Pp 104.

Henry, J, & Henry, Z Doll Play of Pilaga Indian Children New York Amer. Orthopsychiat Assoc, 1944 Pp 133

Sladen, F. J, *Ed* Psychiatry and the War. Springfield, Ill . Thomas, 1943 Pp. 505

White, W Psychology in Living New York Macmillan, 1944 Pp 333

Smith, F T An Experiment in Modifying Attitudes toward the Negro. New York Columbia Univ, 1943. Pp. 135

Deutsch, H Psychology of Women New York: Grune & Stratton, 1944 Pp 399.

Berry, J. R Current Conceptions of Democracy New York Columbia Univ, 1943 Pp 110

Titiev, M Old Oraibi. Cambridge Peabody Museum, 1944 Pp 282

Genetic Psychology Monographs *(continued)*

VOLUME 19—January-December, 1937

1. Psychological bases of self-mutilation—C. Dabrowski
Masculine temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. Gilkinson
3. A psychological study of forty unmarried mothers—R. D. Nottingham
Behavior problems in the children of psychotic and criminal parents—L. Bender
5. Domination and integration in the social behavior of young children in an experimental play situation—H. H. Anderson
6. The sequential patterning of prone progression in the human infant—L. B. Ames

VOLUME 20—January-December, 1938

1. The relationship between characteristics of personality and physique in adolescents—P. S. de Q. Cabot
2. Behavior problems of elementary school: A descriptive and comparative study—I. Young-Masten
Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. Gridley
3. Differences between two groups of adult criminals—R. S. Tolman
4. A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. Troup
Individual differences in the facial expressive behavior of pre-school children: A study by the time-sampling method—C. Swan

VOLUME 21—January-December, 1939

1. An experimental analysis of "level of aspiration"—R. Gould
2. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii—M. E. Smith
3 Domination and social integration in the behavior of kindergarten children and teachers—H. H. Anderson
The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. Galt
4. The social-sex development of children—E. H. Campbell

VOLUME 22—January-December, 1940

1. Measuring human relations: An introduction to the study of the interaction of individuals—E. D. Chapple
2. Aggressive behavior in young children and children's attitudes toward aggression—M. D. Fite
3. Student attitudes toward religion—E. Nelson
The prediction of the outcome-on-furlough of dementia praecox patients—J. S. Jacob
Significant characteristics of pre-school children as located in the Conrad Inventory—K. H. Read
4. Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. McCay, E. B. Waring, and P. J. Kruse
Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. Erikson

VOLUME 23—January-June, 1941

1. An analysis of certain variables in a developmental study of language—F. M. Young
Infant development under conditions of restricted practice and of minimum social stimulation—W. Dennis
An analysis of the mental factors of various age groups from nine to sixty—B. Balinsky
2. Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. Bennett
Individual differences in apperceptive reaction: A study of the response of preschool children to pictures—E. W. Amen

VOLUME 24—July-December, 1941

1. Twins *T* and *C* from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. Gesell and H. Thompson
Finger nail-biting: its incipiency, incidence, and amelioration—A. L. Billig
An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—L. Fromme
2. The Fels child behavior scales—T. W. Richards and M. P. Simons
Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. Smith
Stereotypes in the field of musical eminence—P. R. Farnsworth

VOLUME 25—January-June, 1942

1. A study of factors determining family size in a selected professional group—J. C. Flanagan
A genetic study of geometrical-optical illusions—Sister A. Walters
Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. Read and H. S. Conrad
2. Are there any innate behavior tendencies?—J. B. Schoolland
An investigation of the intelligibility of the speech of the deaf—C. V. Hudgins and F. C. Numbers

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. Miller
Some factors determining handedness in the white rat—K. L. Wentworth
2. Motivation and behavior—E. Frenkel-Brunswik

VOLUME 27—January-June, 1943

1 Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. Maddy
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nine years of age—M. L. Pignatelli

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. Edelston
2. Correlates of vocational preferences—W. A. Bradley, Jr.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. Porteus and R. D. Kepner
2. A twin-controlled experiment on the learning of auxiliary languages—B. Price, W. J. Kostir, and W. M. Taylor

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic apperception test in group situations—R. M. Clark
2. A study of anxiety reactions in young children by means of a projective technique—R. Temple and E. W. Amen

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.